

Kingdom of Saudi Arabia
King Saud University
Deanship of Graduate Studies
English Department

**The Predictive Validity of the Matching Cloze Test as an EFL
Achievement Test for English Majors in the College of Arts
At King Saud University**

By:

Abeer Abdulhadi Hussain

Supervised by:

**Professor Mahmoud Ismail Saleh
Professor of Applied Linguistics**

A thesis submitted in partial fulfillment of the requirements
For the Master's degree in Applied Linguistics

In the Department of English
At the College of Arts,
King Saud University

1424 H/ 2003

:

:

The Predictive Validity of the Matching Cloze Test as
an EFL Achievement Test for English Majors in the College
of Arts at King Saud University

By:

Abeer Abdulhadi Hussain

Thesis Defended: 19 / 4 / 1425 H

M. A. Committee:

Dr. Mahmoud Ismael Saleh

(Chairman)

Dr. Ibrahim Al-Fallay

Dr. Mohammed Al-Alfy

Abstract

The purpose of this study was to investigate the predictive validity of the matching cloze test as a predictive tool for Saudi English majors. The study consists of five chapters. Chapter one is concerned with the introduction, significance of the study, statement of the problem, questions and hypotheses of the study and definition of terms employed in the study.

In Chapter Two, related literature is provided under five sections. The first section explains the significance of integrative testing as opposed to other approaches to testing. The second section deals with the concept of cloze testing from both the theoretical and practical (pedagogical) perspectives. The third section is concerned with the significance and use of several tests and procedures as predictors of future academic success. The fourth section illustrates how the cloze test is used as a predictive tool for success and performance. Finally, the fifth section deals with the use of the cloze test with Arab and Saudi students.

Chapter Three explains the methodology of the study. These include the selection of subjects, instruments for data collection and procedures. Three instruments were used: the matching cloze test, the aptitude test and students' scores on the end of term undergraduate

achievement tests. Finally, at the end of this chapter, a report of a pilot study including 20 students is provided.

Chapter Four presents the analyses of data, findings and interpretations. The data is then subjected to several descriptive and inferential statistical procedures. Such analyses of the results revealed the following: (1) There is a significant correlation between the matching cloze test and the aptitude test, indicating the possibility of using the matching cloze test as an alternative to the aptitude test to predict EFL future success. (2) The matching cloze test scores correlated better with scores of female English majors on their achievement tests than with those of male English majors. (3) The results also showed that the aptitude test is a better predictor of male students EFL future success than it is of female EFL students.

In Chapter Five, conclusions of this study, implications and suggestions for further research are presented. It was concluded that the matching cloze test as well as the aptitude test could be used as valid predictors of future success. Moreover, the researcher favors the use of the matching cloze test over the aptitude test since the former is easier in administration and scoring than the latter. Accordingly, the researcher also recommends using the matching cloze test as an effective testing device for purposes of screening or placement with new admitted EFL students.

10 10

:

(1):

(2) .

(3) .

:

Acknowledgments

I would like to thank several people for their help in writing this research study. First and foremost, I owe a special debt of gratitude to my parents without whose patience, moral support and encouragement the present work would not have been achieved.

I would like to express my deep gratitude to my husband without whose help and continuous support this study could not have been accomplished.

My thanks also go to every member of my family: my children: Noura, Mohammed and Maryam, my sisters and brothers.

I would also like to thank in particular my supervisor, professor D. Mahmoud Ismail Saleh for his constructive guidance and suggestions.

Finally, I would like to thank all the first year undergraduate students enrolled in the English Department of College of Arts for serving as subjects for this study.

Table of Contents

	Page
Abstract	I
Acknowledgments	VII
Table of Contents	VIII
List of Tables	XII
List of Figures	XVI
Chapter One	
1.1 Introduction	1
1.2 Significance of the Study	3
1.3 Statement of the Problem	4
1.4 Research Hypotheses	7
1.5 Delimitation of the Study	8
1.6 Definitions of Terms	9
Chapter Two:	
Review of Related Literature	13
2.1 The Significance of Integrative Testing as opposed to Other Approaches of Testing	13
2.2 The Concept of Cloze Testing	19
2.2.1 The Theoretical Perspective	19
2.2.2 The Empirical and Pedagogical Perspectives	22
2.2.2.1 Studies Focusing on the Appropriate Form and Methodology of the Cloze Test	22

2.2.2.2 Pedagogical Studies Investigating the Advantages of the Cloze Procedure as a Technique for Testing Language Proficiency	28
2.3 The Significance and Use of several Tests and Procedures as Predictors of Future and Academic Success	34
2.3.1 The Use of Aptitude Tests as Predictors	35
2.3.2 The Use of Language-Oriented Standardized Proficiency Tests as Predictors of Academic Success	37
2.3.3 The Use of other Tests and Procedures as Predictors ...	41
2.4 The Use of Cloze Tests as a Predictive Tool of Success and Performance	43
2.5 The Use of the Cloze Test with Arab and Saudi Students ...	46
2.6 Summary	50

Chapter Three:

Methodology

3.1 Introduction	54
3.2 Subjects	54
3.3 Instruments for Data Collection	56
3.3.1 The Cloze Test	57
3.3.2 The Arabic Aptitude Test	60
3.3.3 Scores on Achievement Tests	60
3.4 Procedures	61
3.5 Pilot Study	63
3.5.1 Results and discussion	63
3.5.1.1 Descriptive Testing Characteristics	64
3.5.1.2 Correlation between the Tests	66

Chapter Four:

Presentation and Analysis of the Data

4.1 Introduction	73
4.2 Statistical Procedures Used	73
4.3 Discussion of the Statistical Analysis	75
4.3.1 Several Reliability Coefficient Indices	75
4.3.1.1 The Reliability of Cloze and Aptitude Tests (Cronbach Alpha).....	76
4.3.1.2 The Reliability Coefficients of Cloze and Aptitude Tests (split-half technique).....	77
4.3.2 Descriptive Statistics	78
4.3.3 Inferential Statistics	83
4.3.3.1 t- Test	83
4.3.3.2 ANOVA	85
4.3.3.2.1 ANOVA and Values of Scheffe F-test of Cloze Test: Male Subjects	86
4.3.3.2.2 ANOVA and Values of Scheffe F- test of Aptitude Test: Male Subjects	89
4.3.3.2.3 ANOVA and Values of Scheffe F-test of Cloze Test: Female Subjects	91
4.3.3.2.4 ANOVA and Values of Scheffe F- test of Aptitude Test: Female Subjects	94
4.3.3.3 Pearson Product-Moment Correlation Coefficient r	95
4.3.3.3.1 Correlation between Cloze and Aptitude Tests.....	95
4.3.3.3.2 Correlation Coefficients between Cloze and Achievement Tests: Male Subject.....	97
4.3.3.3.3 Correlation Coefficients between Aptitude and Achievement tests: Male Subjects	99

4.3.3.3.4 Correlation Coefficients between Cloze and Achievement Tests: Female Subjects	101
4.3.3.3.5 Correlation Coefficients between Aptitude and Achievement Tests: Female Subjects	104
4.4 Regression	106
4.4.1 Regression Analysis: Male Students	107
4.4.2 Regression Analysis: Female Students	113
4.5 Discussion of the Results	118
 Chapter Five:	
Conclusions	
5.1 Introduction	126
5.2 Conclusions	126
5.3 Implications of the Study	130
5.4 Suggestions for Further Research	133
 References	 136
 Appendix 1	
Appendix 2	

List of Tables

	page
Table 1 Means and standard deviations of cloze test, aptitude test and achievement tests' scores	64
Table 2 The reliability coefficient for the cloze test and aptitude test	66
Table 3 Correlation coefficients between cloze test results and aptitude test scores	67
Table 4 Correlation coefficients between cloze test results and achievement test scores	69
Table 5 Correlation coefficients between aptitude results and achievement test scores	71
Table 6 The reliability coefficient of the cloze test and aptitude test (Cronbach Alpha)	76
Table 7 The reliability coefficient of the cloze and aptitude tests using split-half procedure	77
Table 8 Means and standard deviations of cloze test, aptitude test and achievement tests' scores	79
Table 9 Means and standard deviations of cloze test, aptitude test and achievement tests' scores (males)	80
Table 10 Means and standard deviations of cloze test, aptitude test and achievement tests' scores (females)	81

Table 11	Male and female subjects' performance on the cloze test / t-Test	83
Table 12	Male and female subjects' performance on the aptitude test/ t-Test	84
Table 13	ANOVA (One-Way Analysis of Variance) Cloze Test: male subjects	87
Table 14	Values of Scheffe F-test of the difference between the male subjects' means on the cloze test	88
Table 15	ANOVA (One- Way Analysis of Variance) Aptitude Test: male subjects	90
Table 16	Values of Scheffe F-test of the difference between the male subjects' means on the aptitude test	91
Table 17	ANOVA (One-Way Analysis of Variance) Cloze Test: female subjects	92
Table 18	Values of Scheffe F-test of the difference between the female subjects' means on the cloze test	93
Table 19	ANOVA (One-Way Analysis of Variance) Aptitude Test: female subjects	94
Table 20	Correlation coefficients between cloze test results and aptitude test scores	96
Table 21	Correlation coefficients between scores of cloze test and scores of achievement tests of male subjects	99

Table 22	Correlation coefficients between aptitude results and achievement tests scores of male subjects	101
Table 23	Correlation coefficients between scores of cloze test and scores of achievement tests of female subjects	103
Table 24	Correlation coefficients between scores of aptitude test and scores of achievement tests of female subjects	105
Table 25	Regression-Model Summary	108
Table 26	Regression coefficients of male students: dependent variable Eng. 120	108
Table 27	Regression- Model Summary	109
Table28	Regression coefficients of male students: dependent variable Eng. 118	109
Table 29	Regression-Model Summay	110
Table 30	Regression coefficients of male students: dependent variable Eng. 113	110
Table 31	Regression-Model Summary	111
Table 32	Regression coefficients of male students: dependent variable Eng. 112	111
Table 33	Regression-Model Summary	112
Table 34	Regression coefficients of male students: dependent variable Eng. 111	112
Table 35	Regression-Model Summary	113

Table 36 Regression coefficients of female students: dependent variable Eng. 120	113
Table 37 Regression-Model Summary	114
Table 38 Regression coefficients of female students: dependent variable Eng. 118	114
Table 39 Regression-Model Summary	115
Table 40 Regression coefficients of female students: dependent variable Eng. 113	115
Table 41 Regression-Model Summary	116
Table 42 Regression coefficients of female students: dependent variable Eng. 112	116
Table 43 Regression-Model Summary	117
Table 44 Regression coefficients of female students: dependent variable Eng. 111	117

List of Figures

	page
Figure 1 Correlation between the cloze test scores and the aptitude test scores	97

Chapter One

1.1 Introduction

Over the past years, many research studies have been conducted to find out the best approaches and methods of language teaching, learning and testing (Oller, 1979; Heaton, 1988; and Cook, 1996). Among the many approaches that have highly governed the teaching and learning of languages, especially of foreign languages, are the structuralist approach and the communicative approach. The structuralist approach views language as composed of discrete elements and skills: elements such as phonology, morphology, syntax and vocabulary, as well as skills such as the skill of listening, speaking, reading and writing. These elements and skills of a language according to this approach can be taught, learnt and tested separately. They are treated as divorced from any context, lacking any pragmatic or communicative purposes.

The second approach that has developed as a reaction to the structuralist approach is the integrative or communicative approach. This approach emphasizes communicative competence as the basic goal to teach and to learn a foreign language. The elements and the skills of a language within this approach are taught, learnt and tested simultaneously, and "they are primarily concerned with meaning, especially meaning in context" (Richards and Rodgers, 1986: 65).

Nowadays, the concept of language as discrete linguistic items is under attack, especially with the decline of the structuralist approach and the audio-lingual method which emphasize repetition and habit formation. Language and language teaching today are best approached from a communicative perspective. This evolution in teaching and learning methods needs a similar change in the way the ability of students should be tested and predicted.

Therefore, this study is mainly concerned with a type of integrative test which does not seek to separate language skills into neat divisions. This type of test is known as the cloze test. A cloze test combines the advantages of other integrative tests, such as the essay writing tests, and the advantages of the discrete test items, such as the multiple choice questions. For example, the cloze test is easy to construct, economical in administration, compared to the discrete linguistic test items, and also objective compared to the subjectivity of the scoring of essay writings.

It is hoped that this research would provide insight into utilizing cloze procedure to predict students' achievements in many EFL undergraduate introductory courses. It also tries to throw light on a specific form on which the cloze test could be based, suitable to predict students' achievement and success in EFL undergraduate courses.

1.2 Significance of the Study

The significance of the study lies in its theoretical and practical implications. Theoretically, it will enlarge our understanding of the concept of integrative testing as well as establish solid grounds for using the cloze test as a valuable procedure to predict students' achievement and success in FL.

On the practical side, the findings of this research will hopefully help the English Department and its teachers in relying on the use of the cloze procedure to predict the outcomes of their students. This prediction may be helpful to the English Department in redesigning some of the programs or the courses to suit students' needs and capabilities. It could also guide the English Department to decide on the proper use of the cloze test as an admission test to select students that join the Department. Moreover, teachers, on basis of this prediction, may challenge their students with more difficult tasks once they have been informed that they have a fast class, or they may redesign easier tasks if facing a slow class.

Similarly, on basis of the procedure of first-day cloze test, students may gain several benefits since it gives them the opportunity to answer an English test and to revise their prior experience and knowledge of EFL. In addition, it would help students to evaluate themselves from the first day they choose to major in EFL and to reflect on their own level of preparation.

1.3 Statement of the Problem

Freshmen coming from high schools have no objective idea of the nature of work and studies that are expected from them during their university studies. This case appears clearly when they choose to major in a foreign language such as English even though it is supposed that these students have acquired a certain level of English proficiency during their schooling. As well, teachers of English Departments have little idea of the capabilities and needs of these new applicants that lead them to succeed in their EFL university courses. Furthermore, English Departments are faced yearly with the problem of over admittance without checking students' potentiality in learning a foreign language.

In an attempt to solve such problems, several tests of English as a foreign language that evaluate students' proficiency and predict their success in learning a foreign language have been developed and validated. However, such tests were criticized from both theoretical and practical perspectives. From a theoretical perspective, the Test of English as a Foreign Language (TOEFL) or the Michigan Test of English Language Proficiency, as examples, mainly evaluate and predict the ability of learning isolated skills and elements. These tests do not pay enough attention to the overall language ability which emphasizes the functional and communicative aspects of language. Practically speaking,

these tests are expensive and time consuming in both administration and construction.

The need for another promising way to provide valid tools for predicting EFL students' achievement and success was evident. This need resulted in the development of some aptitude tests. Several aptitude tests were constructed to assess the ability of students to learn a foreign language. These tests are constructed in the students' native language to measure their ability to learn a foreign language that they have never been exposed to before. An Arabic aptitude test was developed by Al- Khalaf (2000) to predict Arab students' ability to learn English as a foreign language. This test neglected the fact that new admitted students were exposed to English language for not less than six years of schooling. This means that English as a subject in preparatory and secondary schooling is of critical status since the passing and failing of an academic year depends on the grades students earn.

Therefore, there is a need for a simple, valid and inexpensive testing procedure that is capable of evaluating students' gained performance in English as a foreign language, and of predicting students' future success in EFL undergraduate courses.

This study attempts to propose solutions for the major problem mentioned above through investigating the area of the cloze procedure as an example of integrative tests. Stated more specifically, this study is

interested in examining whether the cloze test could be used as a valid and reliable alternative to the aptitude test proposed by Al-Khalaf. Both aim at predicting students' class performance and success in many EFL university courses.

This study aims at accomplishing the following objectives: (1) Investigating the nature of correlation between the subjects' scores on the cloze test used in this study and their scores on their achievement tests administered at the end of the term. (2) Investigating the nature of correlation between the examinees' scores on the proposed aptitude test and their scores on their achievement tests. (3) Investigating the nature of correlation between the cloze test and the aptitude test. These correlations are examined although the major focus of this study is to investigate the correlation between the subjects' scores on the cloze test and their scores on their achievement tests that are administered at the end of the term. This correlation is investigated in order to see if the cloze test could be used as a valid predictor of students' success and achievement in EFL courses. To conduct this study from an empirical perspective, answers to the following major question and related corollary questions are sought:

Major Question:

Can the cloze test be used as a predictive tool of EFL students' class performance and success, as measured and predicted by an Arabic aptitude test?

Corollary Questions:

1. Do the Saudi English majors' scores on the cloze test correlate with their scores on the Arabic aptitude test; both administered at the beginning of an academic year?
2. Do the Saudi English majors' scores on the cloze test correlate with their scores on their future achievement tests?
3. Do the Saudi English majors' scores on the Arabic aptitude test correlate with their scores on their future achievement tests?
4. Is there is a difference between Saudi male English majors and Saudi female English majors with regards to performance on the cloze test and on the aptitude test?
5. Is the cloze test a better predictor of Saudi male English majors' success and future achievement than the aptitude test used for the same purposes?
6. Is the cloze test a better predictor of Saudi female English majors' success and future achievement than the aptitude test used for the same purposes?

1.3 Research Hypotheses

The following null hypotheses will be examined:

1. There is no significant correlation, positive or negative, between EFL students' scores on the matching cloze test and their scores on their end of semester achievement tests.
2. There is no significant correlation, positive or negative, between EFL students' scores on the aptitude test and their scores on their end of semester achievement tests.
3. There is no significant correlation, positive or negative, between EFL students' scores on the matching cloze test and their scores on the aptitude test.

The following null sub-hypotheses will also be examined:

1. There is no significant correlation, positive or negative, between EFL students' scores on the cloze and the same students' scores on the achievement test of Eng.120 (the course of vocabulary building), and there is no significant correlation, positive or negative, between EFL students' scores on the aptitude test and their scores on the achievement test of Eng.120.
2. There is no significant correlation, positive or negative, between EFL students' scores on the cloze and the same students' scores on the achievement test of Eng.118 (the course of translation), and there is no significant correlation, positive or negative, between EFL students'

scores on the aptitude test and their scores on the achievement test of Eng.118.

3. There is no significant correlation, positive or negative, between EFL students' scores on the cloze test and scores of the same students on the achievement test of Eng.113 (the course of reading comprehension), and there is no significant correlation, positive or negative, between EFL students' scores on the aptitude test and their scores on the achievement test of Eng.113.

4. There is no significant correlation, positive or negative, between EFL students' scores on the cloze test and scores of the same students on the achievement test of Eng.112 (the course of listening and speaking), and there is no significant correlation, positive or negative, between EFL students' scores on the aptitude test and their scores on the achievement test of Eng.112.

5. There is no significant correlation, positive or negative, between EFL students' scores on the cloze test and the scores of the same students on the achievement test of Eng.111 (the course of basic language skills), and there is no significant correlation, positive or negative, between EFL students' scores on the aptitude test and their scores on the achievement test of Eng.111.

1.5 Delimitation of the Study

This study is delimited in three ways: First, it considers primarily the predictive validity among the many others that are necessary as characteristics of a good test. Predictive validity means that a test has the potential of predicting future success. Second, the study is limited to one form among the many forms of the cloze test. This form is given with a list of words from which students select the appropriate answer to fill in the blanks of the passage. Therefore, we can refer to this form as a selected response test or as a matching cloze test (Powell and Gillespie, 1990). Thirdly, the study is limited to students in the first academic year at the English Department of King Saud University. Finally, the study is also limited to achievement of first year Saudi English majors in certain English courses.

1.6 Definition of Terms

Achievement test: “An achievement test is related directly to classroom lessons, units, or even a total curriculum within a particular time frame” (Brown, 1994: 259). For the purposes of this study, achievement tests refer to the tests that are administered at the end of first term of EFL university study. They are five achievement tests administered to measure Eng.120 (vocabulary building), Eng.118 (translation), Eng.113 (reading

comprehension), Eng.112 (listening and speaking), and Eng.111 (basic language skills).

Aptitude test: “[It is] the type of test that is given to a person prior to any exposure to the second language, a test that predicts a person’s future success” (Brown, 1994: 259). For the purpose of this study, aptitude test refers to the Arabic aptitude test developed by Al-Khalaf (2000) to predict female Saudi English majors' ability and success in EFL courses.

Cloze test: A cloze test in its basic form is a passage from which every nth word is deleted and in which students are asked to supply the deleted words. It could be based on a passage of general content or of specific content, but relevant to the group of students. It has a variety of formats such as the fixed-ratio deletion and the rational deletion format, or it may have the form of a fixed-response cloze test and the form of an open-ended cloze test. Another way in which the cloze test can be defined is in terms of its scoring method such as the exact word scoring and the acceptable word scoring methods (Cohen, 1980: 91-99).

Exact and acceptable word scoring: Exact word scoring means that “one mark may be awarded for each exact answer”, whereas acceptable word scoring means “one mark may be awarded for each acceptable answer” (Heaton, 1988: 17).

Fixed-ratio deletion and rational deletion methods: “A fixed-ratio deletion method establishes the deletion of every nth word... regardless of what

that word may be, and of the predicted difficulty of the testee “guessing” it” (Brown, 1994: 263).

“A rational deletion provides an alternative to the fixed-ratio method by selecting words for deletion that meet certain discourse criteria” (Brown, 1994: 263).

Fixed-response and open-ended cloze test: A fixed-response test involves “the supplying of alternate response choices for each blank in turn” (Cohen, 1980: 129). An open-ended cloze test is a term used to “refer to those questions which elicit a completely subjective response on the part of the testees” (Heaton, 1988: 133). For the purposes of this study, the cloze test refers to the matching cloze test "constructed from reading passages by blocking off sets of sentences in which five mutually exclusive words from each blocked-off section are deleted and the deleted words are presented in random order in the right margin" (Baldauf, R. and Propst, I., 1979: 3).

Predictive validity: It involves, for example “...[the] use of a test of English as a second language to screen university applicants and then correlate test scores with grades made at the end of the first semester” (Harris, 1969: 20). For the purposes of this study, the predictive validity of the matching cloze test is calculated in regard to the achievement tests administered at the end of first term of EFL university study.

Chapter Two

Review of Related Literature

This section is devoted to reviewing selected literature that deals with testing. It is mainly concerned with previous studies related to cloze testing, so as to derive insights from works done by previous researchers in this specific type which is the area of investigation in this study. The literature is reviewed under the following headings:

1. The significance of integrative testing as opposed to other approaches of testing.
2. The concept of cloze testing from two perspectives: the theoretical and the empirical (pedagogical) perspective.
3. The significance and use of several tests and procedures as predictors of future and academic success.
4. The use of the cloze test as a predictive tool for success and performance.
5. The use of the cloze test with Arab and Saudi students.

2.1 The Significance of Integrative Testing as opposed to Other Approaches of Testing

Some researchers such as Richards and Rodgers and Cook have studied the major distinctive features of several methods and approaches

adopted in language teaching and learning. For example, Richards and Rodgers (1986) tried to explain each teaching and learning approach in a specific chapter throughout their book, whereas Cook (1996) summarized, in the last chapter of his book, the main approaches for teaching and learning the second language. However, these researchers agreed upon the fact that after the decline of the audio-lingual method, based on the structuralist approach, a new approach has evolved. This new approach considers communication as the basic goal of language teaching and learning.

This new approach was introduced by several researchers. Carroll, for instance, introduced it as “the integrative philosophy or approach” (Carroll, 1961, quoted in Oller and Perkins, 1978: 38). This integrative approach necessitates the treatment of language in a more holistic manner in teaching, learning and testing. Its related type of test is also known as the integrative test (Oller and Perkins, 1978).

Oller (1979) distinguished three types of language tests: the discrete point, integrative and pragmatic tests. According to him, discrete point test items attempt to focus attention on one point at a time. Its aim is to test only one element of a language, such as phonology, grammar or vocabulary, or one skill or even one aspect of this element or skill, such as the skill of listening, speaking, reading or writing. Finally, an aspect of an element or a skill could be identified with production or recognition.

An example of a discrete point test item aiming at testing grammar, from the point of view of recognition, may have the form of a multiple choice test item such as:

-Choose the correct suffix to form a noun from the following

adjective: practical

A-...ness

B-...tion

C-...ity

D-...ment

Moreover, Oller classified the discrete point test as based on the structuralist approach, since it tests language divorced from any context. As a reaction to this type of test items, the integrative test was developed with the aim of testing more than one skill at a time. Oller equated integrative to pragmatic communicative tests, since both cause the learner to use the language as in real-life situations. He stated that “integrative tests are often pragmatic in this sense, and pragmatic tests are always integrative” (Oller, 1979: 38). According to him, dictation, cloze tests, as well as composition or essay writings, are all examples of pragmatic or integrative tests. All test the same basic underlying competence with little loss of information in terms of students’ performance.

Oller and Perkins (1980) published some research articles pointing out the importance of integrative tests and how they test the overall

language proficiency of students in an efficient way (see the section on cloze test-based empirical studies). As an example to be given here is the research article entitled “Cloze as an alternative method of ESL placement and proficiency testing” (Hinofotis, 1980: 121). The results of this article indicate that the cloze test, as an example of integrative tests, could be used as a valid and reliable alternative to test students’ performance in place of some discrete point tests.

Another influential researcher interested in testing is Cohen (1980) who distinguished the notions of discrete point, integrative and pragmatic testing but with the belief that there is a continuum from the most discrete point test to the most global or pragmatic test. He explains that a discrete multiple choice item aimed at testing one aspect of grammar, for example, is capable of testing the skill of reading since the student reads and comprehends the question, at the same time, before starting to answer. Therefore, there is no sharp dividing line between the three types of tests.

Spolsky (1986) also explained in his article that there are mainly three conflicting and complementary approaches to the description of language teaching and learning. These are the structural, functional and the overall. Moreover, he tried to differentiate between two important scholars such as Lado and Carroll. According to him, Lado was concerned with relating language to linguistics, emphasizing the point that language

is composed of separate skills which can be broken down into smaller components, permitting in turn the teaching and testing of language components separately. Contrary to Lado was Carroll who was concerned with relating language to psychology and with the total view of language for which integrative and functional tests are very popular. Spolsky referred to the importance and “triumph” of functional communicative tests over the discrete point tests. However, he suggested several solutions to the problems of these tests; for example, he provided several models of scales and of samplings of the materials to be included in functional communicative tests.

Following Oller, Perkins, Cohen, and Spolsky, some other researchers addressed the concept of integrative testing such as Heaton. Heaton (1988) classified language tests in terms of four main approaches of testing: 1) the essay-translation approach, 2) the structuralist approach, 3) the integrative approach and 4) the communicative approach. He defined the integrative approach as involving “the testing of language in context and is thus concerned primarily with meaning and the total communicative effect of discourse” (Heaton, 1988: 16). In this respect, we could notice that his concept of integrative testing did not differ very much from the other previous researchers. However, he classified cloze procedure, dictation, translation and essay writing as integrative tests, and not as pragmatic communicative tests. Moreover, Heaton clarified that

pragmatic tests, based on the communicative approach, emphasize the ability to use language in communication. They test the learner's ability to handle effective communication in the language such as role-plays or writing diaries.

Other researchers interested in identifying differences between the different types of tests such include Powell and Gillespie. They classified tests into two types "constructed-response vs. selected response tests" (Powell and Gillespie, 1990: 3). Constructed-response tests, such as cloze test or essay tests, ask students to construct their own answers whereas selected-response tests, such as multiple choice tests, ask students to select an answer among alternatives. According to them, both types of tests have advantages and disadvantages, and the efficiency of the evaluation of students' performance is totally dependent on the purpose and the good use of the test, not on the type of the test itself.

Norris (2000) examined several types of tests such as the multiple-choice, cloze and computer-based tests. He established that it is the responsibility of the teacher to select, use and evaluate language tests in order to use them in an efficient manner. Moreover, he stated that "a 20-item cloze test, which asks the examinee to write single-word responses to complete a reading passage, provides a very different kind of information than does a 20-item multiple-choice reading comprehension test" (Norris, 2000: 19).

Thus, as can be noted from the literature above, there seems to be a general adoption of the concept of integrative testing, especially with the emphasis and development of communicative-centered approaches of language teaching and learning.

2.2 The Concept of Cloze Testing

2.2.1 The Theoretical Perspective:

From a theoretical point of view, the cloze procedure is reviewed from the perspective of integrative philosophy. Integrative philosophy is the philosophy behind the aims of cloze procedure (Oller and Perkins, 1978).

The researcher who developed the cloze procedure is Taylor (1953), who used it to measure the readability of prose. He defines cloze as “a term Gestalt psychology applies to the human tendency to complete a familiar but not quite finished pattern to see a broken circle as a whole one, for example, by mentally closing the gaps... The same principle applies to language” (Taylor, 1953, cited in Bastidas, 1984: 20). The development of the use of the cloze procedure was shown in Bastidas's (1984) study. He pointed out that the cloze procedure had been used as a testing technique since its inception as early as 1953. It was first used as a testing technique with English as a mother tongue, then with English as a

foreign language. The researcher also referred to the importance of the cloze procedure to test the skill of reading comprehension in particular, explaining the relations between the psycholinguistic principles of reading and cloze procedure. Moreover, he included several ways and steps showing and explaining how the cloze procedure could be used not only to test reading comprehension but also to teach this skill in an efficient manner in the classroom.

Oller (1973) focused on the most widely and successfully used procedures for the cloze test; that is, the measurement of second language proficiency. The cloze test in this case measures the sort of competence revealed in memory constraints, which is an aspect of the underlying competence of second language speakers, since language competence, according to Oller, is best characterized by a grammar of expectancy.

Several researchers described the cloze test as a test in which learners or students are asked to supply words that are systematically deleted from a reading passage, often every n th word such as every 6th, 7th, or 8th word (Oller, 1979; Cohen, 1980 and 1991; Heaton, 1988; Brown, 1994). These deleted words are replaced by blanks of a standard length, and students have to fill in the blanks with the appropriate words. The text should be of a length that would permit a reasonable number of deletions to take place. For example, a cloze test with 50 deletions would

be considered more reliable than a cloze test with 30 deletions (Cohen, 1980; Oller, 1979; Heaton, 1988).

As regards the performance on the cloze test, Cohen (1980), and Heaton (1988) identified three types of knowledge required to perform the cloze test correctly. These are linguistic knowledge, textual knowledge and knowledge of the world. Linguistic knowledge relates to the use of the appropriate lexical item in a grammatically acceptable way. Textual knowledge, as explained by Cohen, involves, for example, “the perception of the cohesive relationship between an item and the rest of the sentence, paragraph, or text as a whole” (Cohen, 1980: 97). Knowledge of the world involves, for example, a reference to a particular century, field, name, date, etc.

Other related points to the cloze test that were stressed by several researchers are the systems of deletion and scoring of the cloze test (See the section on cloze-test-based empirical study focusing on its appropriate form and methodology).

In conclusion, there seems to be a general agreement among these researchers on the advantages and the possibilities in which the cloze test could be used. These could be listed as follows:

1. To measure learners' or students' ability to read a passage.
2. To test global reading comprehension skills.

3. To check learners' or students' awareness of the grammatical aspects of the elements in a passage.
4. As a placement test to determine the relative level of students.
5. As a measure of students' achievements of the materials they have been taught.
6. As a good measure of other language skills such as listening, speaking, and writing (Cohen, 1980 and 1991; Oller, 1979; Heaton, 1988; Valette, 1977).

2.2.2. The Empirical and Pedagogical Perspective:

The cloze procedure from an empirical, pedagogical perspective is reviewed below:

2.2.2.1. Studies Focusing on the Appropriate Form and Methodology of the Cloze Test:

In recent years, there has been a growing interest in pedagogically oriented empirical studies focusing on the cloze test from the perspective of its particular form and from the perspective of its methodology. I mean by the methodology of the cloze test its system of deletion and its method of scoring.

Alderson (1979) in his study that dealt with the cloze procedure and proficiency in English as a foreign language noted that individual cloze tests vary greatly as measures of EFL proficiency. His study was

designed to investigate the effect of certain methodological variables such as deletion rate, text difficulty and scoring procedure on the cloze test. The researcher chose three texts of different levels of difficulty with 650 words in length from the area of fiction. Four deletion rates were used: every 6th, 8th, 10th and 12th word. These deletion rates were applied to the three cloze tests, giving twelve cloze tests, each with 50 deletions. These tests were administered to 360 non-native speakers of English in the UK, and they were scored by five different scoring procedures. The results of the study showed that the cloze procedure is not a unitary technique, since it resulted in tests which are markedly different. These differences were due to differences in deletion rate, text level of difficulty and scoring procedure. Therefore, Alderson advised teachers and test makers to validate and modify each cloze test on its own right to produce the measure they expect. According to him, the cloze procedure is only a technique for producing tests and is not an automatically valid procedure.

Another empirical study was conducted to contrast a fixed-ratio deletion against a selective rational deletion of the same text by Bachman (1985). The subjects of this study were 910 students, including native and nonnative speakers. Half of them took the fixed-ratio test whereas the second, randomly assigned, half took the selective rational cloze test. Bachman concluded that the two tests had equal reliability and validity.

Other researchers discussed the interpretation of cloze test scores and results by identifying factors related to cloze item difficulty. Chapelle and Abraham (1992) in their study investigated the language abilities as measured and indicated by cloze test scores in order to provide a firmer basis for developing various types of cloze tests. For the researchers, test validity is no longer content, concurrent and construct. Instead, they define validity as "a unitary concept which refers to the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores" (Abraham and Chapelle, 1992: 468). The researchers in this study investigated the effect of extrinsic and intrinsic factors on the performance of the cloze test. Intrinsic factors "are the basic cognitive processes required to perform the item task, and the type of response required" whereas extrinsic factors "include students' knowledge of item content and their strategies used to perform the item task" (Abraham and Chapelle, 1992: 492). These factors were examined in three cloze tests (fixed ratio, rational fill-in and rational multiple choice) constructed from a single passage. There were 178 international subjects enrolled at intermediate and advanced ESL composition courses. These main research questions were addressed: (1) how did intrinsic characteristics of items affect item difficulty in the three cloze tests? And (2) how did the extrinsic factors affect performance on the three cloze tests? The results

of the study have implications for cloze test development and use. For example, on the fixed-ratio cloze the interpretation of cloze test scores could be attributed to students' ability to retrieve content words from long term memory, or to their ability to produce words in their correct morphological forms. On the rational cloze items, students' scores could be interpreted as indicating ability to identify context clues and the good use of some psycholinguistic processes such as pragmatic expectancy grammar, enabling students to provide "sophisticated", specific content words and appropriate inflections. On the rational multiple choice cloze test, the interpretations of scores failed to prove constant results. So, researchers recommended further research which may examine the factors mentioned and not mentioned to contribute further to an understanding of an important language testing method such as the cloze testing procedure.

Brown (1993) investigated the characteristics of natural cloze tests. In his study natural cloze tests were defined as cloze procedures developed without interference based on the test developer's knowledge and intuitions about passage difficulty, suitable topics, etc.... The researcher selected fifty reading passages randomly from an American public library. Each passage was made into 30-item cloze test. The subjects were 2298 EFL students from 18 colleges and universities in Japan. The 50 cloze tests were compared in terms of descriptive,

reliability and validity testing characteristics. Brown addressed these three sets of research questions: (1) How difficult are natural cloze tests, and how well do they disperse the scores of students? Are the scores on natural cloze tests normally distributed? (2) To what degree are natural cloze tests reliable as indicated by various internal consistency indices? (3) To what degree are natural cloze tests valid as indicated by correlation coefficients with an outside measure? In what other ways can natural cloze tests be considered valid? The results of this study indicate that a cloze test is not necessarily and automatically a sound overall ESL/EFL proficiency measure. In other words, they indicate that natural cloze tests are not necessarily well -centered, reliable and valid. According to the researcher, tester/researcher intervention is necessary for developing sound cloze tests. Thus, test developers can be helped by the results of the study to create cloze tests that function better in language programs.

After three years, Farhady and Keramati (1996) investigated a text-driven method for the deletion system of a cloze test. They used 403 subjects and 9 versions of the same cloze test. The first version used a fixed-ratio deletion at the rate of 7, and the remaining 8 versions used rational deletion based on the number of the linguistic and discourse structures of the passage. The results of these 9 cloze tests were correlated with the results of a form of a “Comprehensive English Language Test” employed as the criterion measure. The findings of this

study do not support the fixed-ratio random deletion; instead, it favors the semi-nth random deletion rate determined on the basis of text characteristics.

Several researchers, such as Oller, investigated some aspects concerning the methodology of the cloze test. Oller (1972) in his article examined some of the perplexing questions concerning the scoring method, difficulty level of the text used to construct the cloze test and the skills involved in performing a cloze task. The researcher administered three cloze tests of varying levels of difficulty to 398 foreign students. Each of the cloze tests was scored by five different methods: the first method counts the number of exact word provided, the second one allows exact words and any other contextually acceptable response, the other three methods weigh responses in several different ways. The results of the study favor the use of the acceptable word scoring method, regardless of the level of difficulty of the cloze test. Moreover, it was also shown that cloze tests tend to correlate better with tests that require high-level integrative skills such as dictation or reading comprehension.

As regards the best actual form of the cloze test, several studies were conducted. For instance, Bensoussan and Ramraz (1982) conducted an experimental study to contrast several forms of tests such as the fill-in test, the traditional multiple choice tests and the open-ended cloze test. The fill-in test is a modified version of the cloze test in which blanks are

given a number of alternate responses from which students choose the appropriate answer. The statistical analysis of this study showed that all three test formats yield similar results.

Baldauf and Propst (1979) also conducted an empirical study focusing on several forms of the cloze test. These forms were the matching cloze test, the standard cloze, the multiple choice cloze test and the traditional multiple choice tests. The matching cloze tests were constructed by blocking off sets of sentences of a reading passage in which five words are deleted. These deleted 5 words are replaced by blanks and represented in a random order in the right margin. To decrease guessing factors, it is customary to add a sixth word that is randomly selected from the passage and added to the list in the right margin. Following this process, a test may consist of 50 blanks or 10 blocked off sections each containing 5 blanks. The findings of this study indicate that the matching cloze tests could be considered as reliable and valid measures of reading achievement, especially with beginning level ESL students. However, further research is needed to determine the highest level of proficiency at which the matching cloze is appropriate.

2.2.2.2. Pedagogical Studies Investigating the Advantages of the Cloze Procedure as a Technique for Testing Language Proficiency:

A vast part of the literature on cloze testing has investigated the advantages of a cloze test as an indicator of students' proficiency at

different levels and in different disciplines. Jonz (1976), Irvine, Atai and Oller (1974), Ojerinde (1980), Fotos (1991), Soyoung (1996), and Sasaki (2000) have investigated this aspect of cloze testing, for example.

Jonz (1976) conducted a study to explain how an open-ended cloze test could be modified by item analytical and error analytical techniques into a multiple choice cloze test. The researcher carried out this research because he considered the cloze test as a learner-centered teaching and testing device in second language situations. For him the cloze test "is superbly suited to counter-balance the discrete-item test" (Jonz, 1979: 255). To construct the instrument of data collection, the researcher modified an open-ended cloze test into a short 33 multiple choice cloze test. This modification was based on item-analytical techniques. Then, the researcher administered this modified version of the cloze procedure to students who were entering intensive English as a second language program at a large university in the eastern United States. The modified multiple choice cloze test took less than 20 minutes to administer and about one minute to score per test. This m-c cloze was validated with a criterion measure that was administered to the same students the day before the administration of the cloze test. This criterion measure was the intensive English program's established placement examination which took three hours to administer and about thirty minutes to score. The results of the study showed that m-c cloze and the three hour placement

examination appear to measure very similar things. Therefore, the researcher suggested that m-c cloze could be used in an easier way than the placement test examination since the latter is time consuming to administer and to score.

To check the patterns of correlations between the cloze and dictation with each other and the TOEFL, researchers such as Irvine, Atai and Oller (1974) designed a study. One hundred and fifty nine native speakers of Farsi in Iran took the TOEFL that includes five sections, all of them consisting of multiple-choice questions, two dictations and a cloze test. The results of the study revealed high correlations between cloze test and dictation and between cloze test and the TOEFL total. Therefore, it was concluded that the cloze test could be used as a valid measure of EFL proficiency. Moreover, it was also shown that the highest correlation existed between cloze test, dictation, and the listening comprehension sub-section of the TOEFL. According to the researchers, this common aspect that combine cloze test, dictation and listening comprehension is related to the integrative nature of the tasks they require the examinee to perform on them.

Among the earliest studies conducted to find out the validity of the cloze test as a measure of English language proficiency among Nigerian primary school pupils was Ojerinde's (1980). The study was conducted in two consecutive years, 1977 and 1978. The participants were native

speakers of Yoruba and were taught English as a second language. Nine hundred pupils (aged 9) took part in the study in the first year, while 455 of the same pupils (aged 10) took part during the second year. The number was reduced because of financial constraints. The instrument was an English test constructed every year. It consisted of ten subtests that could be given in the following order: spelling, word recognition, listening comprehension, word picture matching, morphology, cloze test, word understanding reasoning, reading comprehension and composition. Comparison between the scores obtained on the various subtests shows that the cloze test had the highest average correlation coefficient. Therefore, it was concluded that the cloze test could be used to measure the English language proficiency of Yoruba-English bilingual children.

Another researcher interested in the efficiency of the cloze procedure as a testing technique was Sandra Fotos. Fotos (1991) conducted a study to investigate the possibility of using a fixed-ratio deleted, exact-word scored cloze test as an integrative measure of EFL language proficiency. The study was carried out in two stages: (1) the selection of the appropriate cloze test based on the principle of pre-testing and on statistical judgments of certain readability formula and (2) the administration of the final cloze test, the essay writing component and the TOEFL. In this research the exact number of subjects was not mentioned although there were certain conditions to meet in order to participate in

this study such as being in the first year of university study. Cloze test scores were correlated with essay writing scores and TOEFL scores. The comparison between the scores on the different tests and their analysis claims that the cloze test may appear to measure different aspects of language proficiency depending on the level of students tested. Thus, with beginning students, the cloze test appears to measure basic skills within the context of the sentence, similar to a discrete-point test. Therefore, the cloze test could be used at this level as a substitute for the TOEFL. But at the intermediate and higher levels, the cloze test has more language proficiency to measure and appears to assess advanced skills. Therefore, it was concluded that a cloze test is useful as a testing technique and can be recommended as a substitute for an essay, for example.

Following Fotos' study, many researchers examined the validity of the cloze test, as an example of integrative tests, with the essay test. For instance, Soyoung (1996) carried out a study aiming at investigating the concurrent validity of the cloze test with the essay test among Korean university students. Subjects were 129 students enrolled in English as a second language courses in Korea. They presented a homogenous group with regard to nationality, language background, age, and educational level. The study included two kinds of material: a cloze test and an essay test on which scores were correlated. The cloze tests were scored twice

by both exact-word method, which counts only the same words in the original passage as correct, and acceptable – word method, which counts any contextually appropriate words as correct as well as the words in the original passage. A significant correlation between the different tests was measured indicating that writing proficiency may be an important factor in cloze performance. The results of this study were classified in terms of implications: theoretical and practical. Theoretically speaking, the results of the current study may be an answer to the question of what underlying factors are involved in cloze test performance. Furthermore, investigation of the factors to influence cloze performance, both linguistic and nonlinguistic, may lend insights into how such factors operate in cloze solutions. Practically speaking, the confirmation of the cloze procedure as a valid language proficiency test, as it has been correlated with essay tests, suggests the use of the procedure as a valid measure of testing writing proficiency of English as a second language.

Examining the influence of cultural schemata on students' cloze test-taking processes was the main purpose of Sasaki's study (2000). A total of sixty Japanese first year university students majoring in English, participated in this study. Their English proficiency level varied from high to high-intermediate with the majority belonging to an intermediate level although the results of t-tests showed that the participants were not significantly different in terms of their English reading ability. These

subjects were divided into two groups according to the type of cloze test (unfamiliar or familiar) they were randomly assigned to take. The first group completed a culturally familiar version of the cloze test whereas the second group completed an unfamiliar version of the cloze test. The results of this study indicated that changing culturally unfamiliar words to more familiar ones in a cloze test had some impact on the participants' test-taking processes, since the first group did better than the second one. Participants of the first group who were given the culturally familiar cloze test did use information beyond the sentence level for solving items. In this respect, the results suggest that the participants' overall comprehension processes are similar to conventional L1 reading processes, and they confirm the claim that cloze tests can measure higher-order processing skills. In conclusion, the researcher suggested the replication of the present study to confirm the results with different types of texts and learners.

2.3. The Significance and Use of several Tests and Procedures as Predictors of Future and Academic Success

Before any instruction begins, teachers prefer to evaluate their students from the first day they meet. They do so in order to measure students' aptitude for learning and ability of using what they have learnt to solve questions and problems in any future communication situation.

Therefore, many tests in a number of research studies were developed and validated especially in terms of their empirical validity that consists of predictive and concurrent validity to help in deciding upon the test to use as a predictive tool.

Three types of tests or procedures would be of concern here. They are as follows:

1. The use of aptitude tests as predictors.
2. The use of language-oriented standardized proficiency tests as predictors.
3. The use of other tests and procedures as predictors.

2.3.1. The Use of Aptitude Tests as Predictors

Aptitude tests were of concern to many researchers, and many definitions were supplied to explain their types and purposes. One of these definitions was given by Brown (1994) who states that it is “the type of test that is given to a person prior to any exposure to the second language, a test that predicts a person’s future success”. Moreover, he points out that “they are designed to measure a person’s capacity or general ability to learn a foreign language and to be successful in that undertaking” (Brown, 1994: 259).

The most popular aptitude tests that are used in the United States are “the Modern Language Aptitude Test” built by Carroll and Sapon and the “Pimsleur Language Aptitude Battery” developed by Pimsleur (Carroll

and Sapon, 1958, and Pimsleur, 1966, both cited in Valette, 1977: 305, Brown, 1994: 259). According to Valette (1977) and Brown (1994), the Modern Language Aptitude Test and the Pimsleur Language Aptitude Battery were designed around some factors such as phonetic discrimination and coding, grammatical sensitivity, memorizing, motivation and logical deductive and inductive learning styles. These tests are built in English and are used with students whose native language is English to measure their abilities to learn a foreign language.

Stansfield (1989) explained the theory on which a foreign language aptitude test could be based. This theory resembles the factors around which the Modern Language Aptitude Test and the Pimsleur Language Aptitude Battery are built. However, according to the researcher there should be a difference between foreign language aptitude and general ability or intelligence. Therefore, the currently used language aptitude tests need more refinement at the levels of test development, data collection and analysis. For this reason, scholars have continued their research to find out the best method, approach and test that could predict students' future academic success.

Al-Khalaf (2000), in her MA thesis, constructed an Arabic aptitude test, adapting the Modern Language Aptitude Test (MLAT), in order to predict Saudi students' future success in learning a foreign language. She used a total of 570 Saudi female students enrolled at the two English

Departments of Girls' Colleges of Education and of Arts to collect her data. Three instruments were used to conduct this correlational study. The Arabic language aptitude test designed by the researcher herself, a standardized English language proficiency test and the subjects' scores on achievement tests for some courses taken the year in which the study was conducted. The results of Al-Khalaf's MA study revealed a significant correlation between the Arabic language aptitude test and future foreign language achievement and proficiency. She reported a positive significant correlation at $P < 0.001$ between the aptitude test and achievement tests for the sample from the College of Education and the sample coming from the College of Arts. These correlation values range from 0.32 to 0.43, respectively. Therefore, it was concluded that the newly constructed instrument could be used as a useful predictor of future foreign language achievement and proficiency of Arab students. However, Al-Khalaf recommends the use of the Arabic aptitude test with learners of foreign language other than English.

2.3.2. The Use of Language-Oriented Standardized Proficiency Tests as Predictors of Academic Success

Despite the fact that empirical evidence supporting the view that future success and achievement could be predicted from the scores of the

TOEFL is constantly growing, it must be mentioned here that current research arguing for the opposite view still exists.

One of the researchers who defended this point was Sokari. Sokari's study (1981), as stated in Al-Salem's thesis, investigated 11 factors as predictors of success of foreign students in two private religious colleges in Northern California. The researcher concluded that TOEFL scores, among the eleven factors, should be disregarded as a predictor of foreign student's success. He added that high school GPA, another factor, may or may not be a good predictor, depending on the home country of the foreign student.

Another study that was conducted in the same field was Odunze's (1982). Odunze in his research said that 54% of his Nigerian subjects did not like taking TOEFL, and thus they did not perform well on it. So, his results were similar to Sokari's in relation to disregarding TOEFL scores as a predictor of success. Thus, he recommended that stronger methods of determining English adequacy should be developed.

Later, Ayres and Quattlebaum (1992) studied the relationship between grade point average and scores on the TOEFL for 60 male and 7 female Asian graduate students of engineering in their research. Low correlations were found. Therefore, the researchers pointed out that TOEFL is not an effective predictor of total GPA, although it measures English communication skills.

Neal (1998) in her study investigated the relationship between graduate grade-point-average (GGPA) and the total score, as well as the scores in each part of the Graduate Record Examinations (GRE). The researcher also investigated the relationship between GGPA and the total score, as well as the scores on each part of the Test of English as a Foreign Language (TOEFL). The subjects of the study consisted of 47 Indian and Chinese international graduate students. All of the students included in the study completed the GRE and TOEFL tests prior to their admission to prepare a master's degree in science and engineering at Rose-Hulman Institute of Technology. The results of the study showed no significant correlation with the total TOEFL score or any individual TOEFL score and the GGPA. Therefore, the researcher concluded that for international Indian and Chinese students, the TOEFL appears to have little, if any, predictive value regarding GGPA. The correlational study between Graduate Record Examination (GRE) and graduate grade-point-average GGPA also showed no significant correlation between them. Therefore, the researcher discarded the TOEFL and the GRE exams as tools for predicting success in graduate school.

Al-Musawi and Al-Ansari (1999) examined the multivariate relationship of the TOEFL and the First Certificate of English (FCE) to determine the possibility of using students' total score on the TOEFL or their overall score on the FCE as good predictors of their success at the

University of Bahrain as measured by overall GPA. Subjects were 86 students enrolled at the University of Bahrain. The multivariate prediction of GPA from the scores on the FCE was very accurate. Regression analyses revealed that the FCE Cloze and sentence Transformation sub-scores were the only test scores that contributed to the prediction of students' GPA. Therefore, the use of the TOEFL as an effective predictor of students' future academic achievement at the University of Bahrain should be reconsidered.

Sharon (1971) conducted another research in which it was hypothesized that TOEFL would moderate the relationship between the Graduate Record Examination (GRE) Aptitude Tests scores and GPA. According to this hypothesis, students scoring high on TOEFL would be more predictable by GRE than those scoring low. The results of the study suggested that foreign students with low English verbal aptitude can succeed in American graduate schools. Therefore, the hypothesis was not fully rejected, and thus the TOEFL can not be used as a valid predictor of GRE or of GPA.

As a reaction to the claim that TOEFL cannot be used as a predictive tool of future success, Ayers and Peters conducted their study. Ayers and Peters (1977) examined the validity of the TOEFL to predict success of fifty Asian students in engineering, chemistry and

mathematics. Their results showed significant correlation between the TOEFL and grade point average.

As a final example of predictive validity studies regarding TOEFL, to be mentioned here, is the study done by Bostic. Bostic (1981) reached significant correlation in her study that was carried out to investigate the predictive validity of TOEFL as it is used in Oklahoma colleges and universities as an admission procedure. She found that the students' performance in nonverbal fields, scientific or technological-oriented ones, were more easily predicted by the TOEFL scores than in language-oriented fields.

2.3.3. The Use of other Tests and Procedures as Predictors.

Other tests and procedures such as the placement and admission tests are administered to predict future success. These tests may predict general success but not actual grades as well as they can help identify students who need additional assistance.

Many schools administer some type of standardized test to prospective students in an attempt to assess students' potential success. The College of the Canyons (COC) in California (1994) conducted several predictive validity studies of the College Board Assessment and Placement Services (APS) Reading Test. Fall 1993 scores were used in determining the ability of the APS Reading Test to predict student

success in both skills and college-level English courses offered at COC. Correlation coefficients were computed for each English course between the test scores and course grades with withdrawals converted to fails, and between test scores and course grades with withdrawals deleted. The predictive validity value across all levels was .27, indicating a moderate linear relationship between APS Reading Test scores and course grades. However, COC considered APS Reading Test adequate to justify its use for placing students into reading courses at COC.

In a more recent study, Ammeraal (1997) examined the correlation between students' placement test scores on a multiple choice test and their passing rate on the Advanced Placement (AP) language exam. It was hypothesized in the study that there would be a correlation between placement test scores based on the first 50 questions in the 1991 Acorn multiple choice test and passing rate on the 1992, 1995, and 1996 AP language and composition examinations. Sample for the study included 73 students coming from middle to high economic neighborhoods and who were studying high school at a Catholic boys' school. Results indicated that the Acorn multiple-choice scores can be used as predictors of student success on the AP tests. The higher the score on the pretest, the more likely it is that a student would pass the AP exam. However, the researcher pointed out that Acorn scores can predict general success but not actual grades of the AP test.

2.4. The Use of the Cloze Test as a Predictive Tool for Success and Performance

A very important dimension in which the cloze test could be used is for predictive purposes. Several research studies have been conducted to examine the potential of the cloze test to predict future success at different levels and in different disciplines. Hisama, Lewis and Woehlke (1977), Allen (1980), and Stephenes et al. (1986) for example, conducted some of these studies.

One of the studies that proved that the cloze test is highly reliable, valid and efficient, with respect to testing time and ease of scoring is Hisama, Lewis and Woehlke's (1977). This research was carried out to investigate the use of the cloze test as a placement test for foreign students to predict future academic success. One form of the cloze test along with three placement tests were administered to 136 subjects at the Center of English as a Second Language, Southern Illinois University, all recent immigrants coming from different nationalities. Later, at the end of the session, another form of the cloze test was administered along with the Michigan Test of English Language Proficiency. Several forms of correlation between these different tests were calculated, and it was found that the cloze test could be used as a reliable, valid and efficient predictive tool of future performance and could measure ESL proficiency.

This new test satisfies all conditions of a good test, although a few technical problems should be investigated in further research.

Hanzeli (1977) in his study described an experiment undertaken at the University of Washington, involving 107 learners of French on different levels. A cloze test with every 7th word deleted was administered to the subjects during their final examination. Other variables included in the experiment were University of Washington grade point average, final grade in the course in which the test was taken, and length of prior study of French both in college and high school. The results of the study have several implications. First, at the level of scoring the cloze procedure, both acceptable word scoring and exact word scoring rank the student in the same way. Second, scores on the cloze test show high correlation with the students' prior exposure to French and with course grades. Therefore, according to the researcher, cloze scores predicted course grades and could be rated by non-native speakers of the foreign language since both scoring methods are highly correlated.

Another researcher interested in investigating the correlation of first day cloze with final grade in literature for foreign students is Allen. Allen (1980) conducted an empirical study to investigate the above-mentioned correlation. He administered several cloze tests of different fixed deletion rate to 14 classes with a total of 384 students. The results of the study indicated that the cloze test could be used as a predictor of

future accomplishment. Moreover, the researcher states that “these teacher-constructed cloze tests prove to be more accurate than other instruments” (Allen, 1980: 63).

Stephens et al. (1986) conducted a study to examine the adaptability of the cloze procedure as a predictive tool for undergraduate classes. In the study, the relation between cloze performance and student achievement was investigated. There were over 5,000 students from a state college coming from different disciplines such as politics, sociology, psychology and biology in this investigation. Several cloze tests were administered to the subjects and Pearson correlation coefficients were calculated between cloze test scores and students' scores on their achievement tests. In all cases, the coefficients between the results on the cloze tests and the students' final grades were positive, varying from .318 to .590. These results may indicate the possibility of the use of the cloze procedure to predict students' future achievement. However, the researcher recommended the extension of the investigation to other college levels and to other disciplines and subjects in order to throw light on more questions about the cloze test.

2.5. The Use of the Cloze Test with Arab and Saudi Students

Some researchers investigated the notion of the cloze procedure in an Arabian setting. For example, Stubbs & Tucker (1974), Hanania and

Shikhani (1986), Al-Fallay (1994; 1997a and 1997b), and Al-Salem (1998) conducted empirical studies to investigate the cloze test with Arab students.

Stubbs & Tucker (1974) in their study tried to present some empirical data concerning the use of the cloze test as a measure of proficiency in English as a second language. 155 native speakers of Arabic were the subjects of the study. Researchers scored the test twice for each candidate, first for exact responses, and the second for any contextually-acceptable response. They were interested in the relationship which existed between scoring the tests for exact replacement and for contextually-appropriate responses. The results of the study revealed that there was a positive correlation between scoring for exact versus contextually-appropriate responses. This result suggests that the cloze technique may well be appropriate for use by non-native teachers. Regarding the validity of the cloze test, and on the basis of their results, it was concluded that the use of the cloze test constitutes a powerful and economical measure of English-language proficiency for non-native speakers as well as a useful diagnostic tool for the classroom teacher.

One of the studies that investigated the cloze test in an Arabian setting is Hanania and Shikhani's (1986). They investigated the interrelationships among standardized ESL, cloze and writing tests. The study was mainly conducted to determine whether the addition of a cloze

component to the standardized ESL test would improve the predictability of students' communicative proficiency as reflected in their performance on a writing test. The three tests were administered to 1,572 students in the American University of Beirut. The results of the study showed an interesting pattern of interrelationships among the three types of tests. Moreover, the combination of any two of the three tests would improve the predictability of the third, but the cloze and writing tests appeared to measure in common some aspects of language ability beyond those measured by the standardized structural test.

More studies were carried out to examine the use of cloze tests with Arab students. Al-Fallay (1994) in his Ph.D. thesis investigated the role of background knowledge on students' performance through the use of the cloze procedure. The subjects in this study were 74 Saudi male students learning English as a foreign language at the Institute of Public Administration in Riyadh coming from different majors. The cloze tests in this study were based on two texts. The first text was referred to as the Saudi text consisting of five episodes written by the researcher himself whereas the second text was referred to as the American text adopted from EFL reading materials was also divided into five parts. From each text two versions of the cloze test were made: an original and a modified version. Thus, four cloze tests were constructed for the purpose of the study that were based on the original Saudi text, the modified Saudi text

that conforms to the western culture, the original American text and a modified American text that conforms to the Arabic culture. The subjects were divided randomly into two groups: the experimental group and the control group. The experimental group took the cloze test based on the original Saudi text and the cloze test based on the modified American text that agrees with the Saudi background knowledge. The control group took the cloze test based on the original American text and the cloze test based on the modified Saudi text that agrees with the American expectations. The results of this experimental study indicate that when subjects are given tests conforming to their background knowledge they may perform better than those given tests of other background knowledge.

The same researcher evaluated the reliability and validity of the fixed ratio multiple choice cloze test in order to use it as the sole testing technique in assessing EFL Arab students' achievement. Al-Fallay (1997a) conducted a study in this respect. There were 348 subjects coming from the Institute of Public Administration and from four different levels of English proficiency. Although, the results of the study showed that the fixed ratio multiple choice cloze tests could be considered reliable and valid when used as the only technique to assess EFL students' achievements, it is more advisable to incorporate it with other achievement tests. Moreover, the researcher claims that the fixed

ratio multiple choice cloze test and the open ended cloze test should be considered as two autonomous tests. Therefore, the findings of the research that apply to the cloze test should not similarly be reapplied to the fixed ratio multiple choice cloze test.

A third study was also conducted by Al-Fallay (1997b) to examine the relationship between the cloze test and tests of actual courses given at the Institute of Public Administration. 432 Saudi Arabian students enrolled in private sector programs of the Institute of Public Administration were involved in the study. Four cloze tests of fixed deletion rate were administered to these subjects coming from four different levels of English language proficiency. These cloze tests were administered one week before the final examination since the main purpose of the study was to investigate the reliability and validity of the cloze test as an achievement test. The results of the study showed strong relationship between these cloze tests and the different achievement tests that evaluated English language skills separately. Therefore, it was concluded that the cloze test is a promising tool to evaluate students' achievement.

Al-Salem (1998), in her MA thesis, conducted an empirical study to investigate the predictive validity of the English Admission Test Battery in the Department of European languages and Translation of King Saud University. There were 170 female students in this study. The cloze

test constitutes a sub-test among the other tests. The conclusion of her study indicates that a good constructed cloze test may serve as a useful predictor of students' second language proficiency.

2.6 Summary

In the preceding sections, we have reviewed selected literature related to language testing. We have seen that the approach of communicative or integrative testing is more beneficial in most of the cases than that based on the structural or discrete item testing (Richards and Rodgers, 1986; Cook, 1996; Oller, 1979; Oller and Perkins, 1980; Cohen, 1980; Spolsky, 1986; Heaton, 1988; Powell and Gillespie, 1990; Norris, 2000).

Consequently, as a positive reaction to the growing interest in communicative and integrative approach, the cloze test as an example of integrative tests, has been the focus of much research. Cloze procedure initially introduced by Taylor in 1953, was used to determine the readability of texts in the reader's native language. Later, it gained weight as a testing device in native language settings, and now it is popular as a testing technique for assessing language proficiency of ESL learners (Jonz, 1976; Irvine, Atai and Oller, 1974; Ojerinde, 1980; Fotos, 1991; Soyoung, 1996 and last but not least Sasaki, 2000).

As a result of the promising findings of the above mentioned research, much attention has been directed towards factors that may affect and contribute to the effectiveness of the cloze procedure as an EFL proficiency measure. These factors such as scoring methods, deletion system and rate, passage level of difficulty and test length were the purpose of much research in order to investigate reliability and validity estimates of the appropriate cloze test (Alderson, 1979; Bachman, 1985; Chapelle and Abraham, 1992; Brown, 1993; Farhady and Keramati, 1996; Oller, 1972; Bensoussan and Ramraz, 1982; Baldauf and Propst, 1979).

More specifically, for the purpose of the present study, the cloze test was reviewed as a predictive tool. Its use as a predictor was compared to other procedures such as aptitude tests and some language-oriented standardized proficiency tests. Surprisingly, in several studies the cloze test showed its advantage over other existing testing techniques to predict future success (Hisama, Lewis and Woehlke, 1977; Hanzeli, 1977; Allen, 1980; Stephens et al., 1986).

Within the above context, more research was also conducted to investigate the potential of the cloze test as a measure of EFL proficiency with the concern here with Arab and Saudi students (Stubbs and Tucker, 1974; Hanania and Shikhani, 1986; Al-Fallay, 1994; 1997a; 1997b; Al-Salem, 1998).

From the foregoing review, it is clear that cloze testing has been applied to a wide range of purposes. However, none of the studies has attempted to apply an analysis of the cloze test in terms of its forms to first year college ESL students to predict future achievement and performance. The studies listed above of Al-Fallay and Al-Salem addressed of course the cloze test but from different points of view. Al-Fallay (1994) investigated how the background knowledge affects students' performance through the use of different open-ended cloze tests. In 1997, he examined the reliability and validity of the fixed ratio multiple choice cloze test in order to be used as the sole assessment technique to evaluate EFL students' achievement. Finally, in 1997, Al-Fallay analyzed the relationship between the open ended cloze test and the EFL achievement tests which tested English language skills separately. Although, Al-Fallay (1997b) mentioned the possibility of using the cloze test as an aptitude test, it was clear that in his study the cloze test was mainly examined and evaluated as an achievement test, not as an aptitude test. In Al-Salem's study the cloze test constituted only a subtest among other types of the Admission Test.

The present study is conducted to investigate the potentiality of the cloze test as a predictive tool or as an aptitude test that is capable of evaluating students' gained performance in English as a foreign language, and of predicting students' future success in EFL undergraduate courses.

In this way, the current study will make a contribution to the area of the cloze procedure as the study is expected to be based on insights derived from the literature and data obtained from the matching cloze test and the Arabic aptitude test that were administered to freshmen level students.

Chapter Three

Methodology

3.1 Introduction

This chapter is devoted to the explication of the development of the study itself. It includes a discussion of the subjects and the instruments and procedures used to achieve goals of the study. A discussion of a pilot study of 20 subjects is given at the end of this chapter.

3.2 Subjects

This investigation of the cloze procedure was conducted on male and female English majors enrolled at the Department of English at King Saud University during the academic year 1422- 1423 H (2001-2002). They were all new students registered to study EFL and they are required to study five EFL courses during their first term of university study. These are Eng. 111, Eng. 112, Eng. 113, Eng. 118 and Eng. 120.

A total of 202 subjects participated in the study. This number of subjects could be seen as one group because of homogeneity of conditions, or it could be divided into two sub groups: male and female EFL students.

A questionnaire eliciting information about the subjects' social and educational background was administered with the cloze test. On the

basis of this information, a description of the two groups of subjects could be provided.

The female EFL students consisted of 108 students, 33 (30, 5 %) of them transferred from other departments, but they were first year students in the English Department. The ages of the female subjects ranged from 17 to 21 years. As for their educational background, 66 (61, 1%) of them came from public schools, 24 (22, 2%) of them came from private schools and 18 (16, 6%) of them studied at both private and public schools. Fifty two (48, 1%) of the female subjects were specialized in the scientific branch in the secondary school, fifty five (50, 9%) of them had specialized in the literary branch, and the remaining one (0, 9%) had studied the secondary school abroad where both branches are intermixed. Seventy nine (73, 1%) of the female subjects had not received extra training in the English language other than the school curriculum, whereas twenty nine (26, 8%) of them had received extra training. Regarding the grade point average obtained by the subjects at the end of their secondary school, they are as follows. Thirty six (33, 3%) of the subjects had obtained a GPA above 95%, twenty five (23, 1%) had received above 90%, thirty three (30, 5%) of them got a GPA above 85%, nine (8, 3%) of them obtained a GPA above 80%, two (1, 8%) gained a GPA above 75% and the remaining three students did not give their GPA.

The second group consisted of 94 male EFL students. Twenty eight (29, 7%) of them transferred from other departments, but they were at their first level in EFL courses. Their ages ranged from 18 to 22 years. Seventy six (80, 8%) of the male subjects studied at public schools, eight (8, 5%) of them studied at private schools and ten (10, 6%) divided their studies between private and public schools. Sixty three (67, 0%) of the male subjects were specialized in the scientific branch in the secondary school and the remaining thirty one (32, 9%) had specialized in the literary section. Fifteen (15, 9%) of the male students had received extra training in the English language whereas seventy nine (84, 0%) of them had not received any extra training. Six (6, 3%) of the male students got a GPA equal to or above 95%, twenty three (24, 4%) of them gained a GPA equal to or above to 90%, fifty one (54, 2%) received a GPA above or equal to 85%, ten (10, 6%) gained a GPA above 80%, three (3, 1%) of them got a GPA above or equal to 75%, and the last one did not mention his GPA.

3.3 Instruments for Data Collection

This study used three main instruments: 1) a cloze test, 2) an Arabic aptitude test and 3) the subjects' scores on the achievement tests administered at the end of the first term. Here are the descriptions of these instruments.

3.3.1 The Matching Cloze Test

A matching cloze test was constructed on the basis of the form of a selected-response test or what may be called a matching cloze test. This matching cloze test was given to the subjects with an alphabetically arranged list of words from which the students should select words to fill in the blanks of the passage. Another way in which this cloze test could be described relates to its system of deletion. To construct the cloze test used for the study a fixed-ratio deletion of items was adopted. Every 7th word was deleted, but the first and last sentences of the passage were left intact to provide context, and there are 30 deletions. The passage on which the cloze test was based was divided into 7 sections. Each section contained a number of deleted words and these deleted words were provided in a box opposite to each section and were arranged alphabetically. From the list of words of each section, students were asked to fill in the blanks of the passage. In order to decrease guessing factors, an extra word was added to the test. The directions used to construct the matching cloze test were based on Baldauf and Propst (1979) recommendations. A copy of the test is given in the appendix.

Another important point that could be mentioned here relates to the choice of the passage used to construct the cloze test. First, the readability level of first year EFL students was decided by checking the level of

difficulty of two passages extracted from students' English books taught at the secondary level. The readability formulas used to determine the readability level of the passages were: 1) the Flesch Reading Ease readability and 2) the Flesch-Kincaid readability that were explained in Flesch (1948) and Schulz (1981) studies. Here is a detailed explanation of the two passages: first, the first passage contained 291 words, 22 sentences and has a Flesch Reading Ease of 50.1 and a Flesch-Kincaid Grade level of 8. Second, the second passage contained 300 words, 29 sentences and has a Flesch Reading Ease of 68 and a Flesch-Kincaid level of 6.8. On the basis of these calculations of the level of difficulty of these two passages, the choice of the passage used to construct the cloze test was made. The selected passage chosen to construct the cloze test has the following characteristics: 260 words, 6 paragraphs and 22 sentences. It has a Flesch Reading Ease readability of 46, 2 and a Flesch-Kincaid level of 9.5. The passage was taken from page 158 of Kingfisher (2001) and it could be described as authentic, since it is designed to be read by native speakers. Therefore, the choice of the passage for the cloze test of the study could be considered suitable, because it represents an approximate level of difficulty, considering that of the two passages that were extracted from students' English books taught at public schools in the secondary level. The passage could also be described as 'fairly difficult' passage according to the Flesch scale. This choice of fairly difficult

passage is based upon the empirical investigations of some researchers such as Alderson and Oller. According to Alderson (1979) and Oller (1972), difficult tests result in better correlation with proficiency and criterion measures such as some aptitude tests. Moreover, in order to increase face validity, I showed the passage in its cloze form to my professors in the Department and to my colleagues who are preparing their MA thesis in the same field. The cloze test was considered suitable to first year EFL students to be completed in no more than 30 minutes.

Finally, at the beginning of this study, I administered the cloze test in two forms, the constructed-response cloze test and the selected-response cloze test, to ten subjects. The students who were given the constructed-response cloze test could fill only less than the half of the items deleted, so I excluded the use of this form in the study and I only relied upon the selected-response test. This latter form was proven to be more favorable because it can be scored by using the exact word scoring method, and in this way objectivity is assured and time needed for scoring is shortened.

In addition, a questionnaire eliciting educational and social background of the subjects was administered. This questionnaire was in Arabic and it was given to elicit the following information: 1) Name and ID number 2) Age 3) College level 4) Date of entrance to the Department 5) Educational information about the subjects' school attended 6)

Educational information about the subjects' specialization in the third year of secondary school, their GPA and date on which high school certificate was obtained and 7) whether students had received any additional training in English other than the English curriculum taught at school. If so, students were asked to supply place and duration.

3.3.2 The Arabic Aptitude Test

The Arabic aptitude test used in this study was developed by Al-Khalaf. Al-Khalaf (2000) conducted a research study with the goal of developing an Arabic aptitude test capable of predicting students' future success to acquire a foreign language such as English. This Arabic aptitude test consists of three parts: Part I: Spelling Clues "

", Part II: Words in Sentences " " and Part III: Paired Associates " ". For details of the whole test, see Al-Khalaf (2000). A psychometric profile of the test is provided in appendix (2).

3.3.3 Scores on Achievement Tests

Students' scores on their cloze test were correlated to their scores on their five achievement tests administered at the end of the first term of first year EFL university study. Similarly, students' scores on the aptitude test were correlated with scores of the same students on their five achievement tests administered at the end of the first term of first year

EFL university study. These five achievement tests were given for the following courses: Eng.111, Eng.112, Eng.113, Eng.118, and Eng.120. Eng.111 stands for the course of basic language skills, Eng.112 teaches listening and speaking, Eng.113 teaches reading comprehension, Eng.118 teaches translation, and Eng.120 tackles vocabulary building. They are all taught 2 hours per week except for Eng.111 which is taught 3 hours per week.

3.4 Procedures

The administration of the cloze test took place at the beginning of the first and second terms of the year 1422/1423 H. More specifically, it was given to first level EFL King Saud University students during the second week after the beginning of the college term. The present researcher herself conducted the administration of the tests to female students whereas the help of a faculty member at KSU at the male section was vital to administer the tests to male subjects. Both types of tests were administered during normal ESL courses. They were given in classrooms, and the atmosphere in which the subjects performed the tests could be compared to that used in normal or formal tests. The time allowed to finish the cloze test was 35 minutes, five minutes to fill in the demographic information attached on a separate sheet before the start of the cloze test and thirty minutes to fill in the blanks of the cloze passage.

Before the start of the cloze test students were given specific instructions on how to proceed with the cloze task. After the collection of the cloze test, it was scored by the exact word scoring method, since its form is based on the principle of selected-response test.

A week after the administration of the cloze test, the Arabic aptitude test was given to the same students who took the cloze test. But with this test the time allowed to finish it was longer since it consisted of more items. Therefore, it was necessary to administer the test during two class sessions of regular EFL courses. More specifically, this test took 70 minutes that were distributed as follows: a total of fifteen minutes for reading the instructions (five minutes preceding each section), twenty five minutes for part I, twenty five minutes for part II, and five minutes for part III. Parts I and III were completed during a class session whereas Part II was finished the next session. The Arabic aptitude test was scored according to the answer key sheet provided by Al-Khalaf who developed the test. During the administration of Part III, students were carefully monitored to assure their compliance with the direction not to look back at the list of words once they had begun answering the 25 items contained in that Part.

3.5 Pilot Study

In the first portion of the statistical analysis, a pilot study was conducted to check if the cloze test is satisfactorily used as a predictive tool for the final results of the subjects. For the purposes of the pilot study and before administering the test to the total number of subjects, the cloze and aptitude tests were administered to twenty first year university students. Ten were males and ten were females.

The first step in conducting this subsection of the research and after collecting the data was to calculate the descriptive statistics of each of the measures involved in the study. These descriptive statistics included the number of items, subjects, means and standard deviations. The purpose of doing such analysis was to see whether the cloze test used for the present study had sound descriptive statistics characteristics. This analysis helped also to check that the passage selected for the construction of the cloze test was of reasonable difficulty and that there was no need for any modification or refinement regarding the selection of the passage.

3.5.1 Results and Discussion

For the purpose of the pilot study, statistical analyses were performed in many ways: (1) descriptive statistics of the tests, such as means, standard deviations, minimum and maximum scores are presented, (2) reliability estimates of the cloze and aptitude tests using the reliability coefficient of Cronbach Alpha and (3) the Pearson's product-moment

correlation to see how the cloze and aptitude tests correlated with each other. Moreover, it was performed between the cloze and achievement tests to determine the predictive potentiality of the cloze test. As a final step, the same correlational analysis was also performed between the aptitude and achievement tests to see how well the aptitude test could predict students' future success in EFL courses.

3.5.1.1 Descriptive Testing Characteristics

Table (1) displays - for both males and females – a summary of mean, standard deviation and minimum and maximum scores of the subjects (N= 20) on the cloze test, aptitude test, and the various achievement tests.

Table (1)
Means and standard deviations of cloze test, aptitude test and
achievement tests' scores

Test	Descriptive statistics				
	Mean	Standard Deviation	Minimum	Maximum	No.
Cloze test	20.05	5.05	12	29	20
Aptitude test	67.35	21.74	24	100	20
Achievement test 1 Eng.120, Vocabulary Building	57.65	25.10	0	98	20
Achievement test 2 Eng.118, Translation	53.65	28.60	0	97	20
Achievement test 3 Eng.113 Reading Comprehension	59.30	25.99	0	98	20
Achievement test 4 Eng.112 Listening & Speaking	68.20	19.35	10	98	20
Achievement test 5 Eng.111 Basic Language Skills	63.50	22.94	15	90	20

* The maximum score of the cloze test was 30.

* *The maximum score of the aptitude test was 120.

As can be seen in table (1), the mean scores on the cloze test is 20.05 (s.d. = 5.05), on the aptitude test, the mean score is 67.35 (s.d. =21.74). On the achievement tests, the mean score ranged from 53.65 to 68.20. Scores on the cloze test ranged from 12 to 29, and scores on the aptitude

test ranged from 24 to 100. The difference of means and standard deviations between male and female groups would be calculated with the total number of subjects. The means and standard deviations of achievement tests, as shown in table 1, suggests slight differences among the students. This could be related to the point that these students were about equal in EFL proficiency because these EFL achievement tests were administered after one term of academic learning.

The following table shows the reliability coefficients Alpha computed for the cloze test and the aptitude test.

Table (2)
The Reliability coefficients for cloze test and aptitude test

Test	Reliability Coefficients		
	Male Students	Female Students	Both
Cloze test	.8344	.8092	.8076
Aptitude test	.9470	.8770	.9523
No. of Subjects	10	10	20

The results in table (2) show that the reliability estimates were very high for both the cloze and the aptitude test. Therefore, this encouraging result, despite being at the level of a pilot study, indicates that both tests have met an important criterion to be good tests in that they are reliable. One could notice that the reliability on the aptitude test was higher for

both groups, male and female groups. But this could be attributed to the length of the aptitude test which contained 120 items whereas the cloze test contained only 30 items.

3.5.1.2 Correlation between the Tests

For the estimation of the concurrent validity of a test Lado (1961), Valette (1967) and Harris (1969) recommend a method consisting of correlating its results with the results of another test or a battery of tests, of which the validity has been evaluated and confirmed previously. On this basis, the higher the correlation coefficient would be, the more valid the test will be.

Based on the predictive validity approach, this study focuses on assessing the nature of the association between the performance on the cloze test and the aptitude test, between the performance on the cloze test and the student's achievement tests administered at the end of the term, and then between the performance on the aptitude test and students' achievement tests. Pearson correlation coefficient was calculated for the relationship between students' scores on the cloze test and their scores on the aptitude test, between the students' scores on the cloze test and their scores on the achievement tests in five different English major courses and between students' scores on the aptitude test and their scores on the same achievement tests.

Table (3)

Correlation coefficients between cloze test results and aptitude test scores

		Cloze T.
Aptitude T.	Pearson Correlation	.541*
	Sig. (2-tailed)	.014
	N	20

* Correlation is significant at less than the 0.05 level (2-tailed)

** Correlation is significant at less than the 0.01 level

Table (3) presents the relationship between the cloze scores and aptitude scores, which was expressed by the Pearson's product-moment correlation coefficients. The .54 correlation coefficient indicates a degree of commonality between the tests since it is close to the 0.05 level, the level of significant correlation. Therefore, the predictive validity of the cloze test could be confirmed as it was done in previous research with the aptitude test.

Table (4) below presents the second correlation coefficient which is computed between cloze test results and achievement tests scores in the different English courses.

Table (4)
Correlation coefficients between cloze test results and achievement
tests scores

		Cloze T.
Eng.120 T	Pearson Correlation	.582**
Vocabulary	Sig. (2-tailed)	.007
building	N	20
Eng.118T	Pearson Correlation	.667**
Translation	Sig. (2-tailed)	.001
	N	20
Eng.113T	Pearson Correlation	.709**
Reading	Sig. (2-tailed)	.000
Comprehension	N	20
Eng.112T	Pearson Correlation	.323
Listening &	Sig. (2-tailed)	.164
Speaking	N	20
Eng.111T	Pearson Correlation	.595**
Basic	Sig. (2-tailed)	.006
Language Skills	N	20

** Correlation is significant at less than the 0.01 level.

* Correlation is significant at less than the 0.05 level (2- tailed)

As it can be seen from the table, the correlation coefficients range from 0.3 - the correlation value between the cloze test and the English course 112 - to .70 - the correlation coefficient between the cloze test and the English course 113. Hence, table (4) shows clearly that there is a

significant positive correlation between the cloze test and only four achievement tests since the correlation between the cloze test and the English course 112 is 0.3, which is not significant. Therefore, it could be prejudged here that the cloze test could be valid to predict future success of these courses and not the EFL course of 112. This result agrees with that of Al-Fallay (1997) in which the cloze testing procedure has been proved to be as a promising evaluation tool that can be used in a continuum fashion. However, these results would be further analyzed with the total number of subjects' scores.

Following is the relationship between the aptitude test and the five achievement tests administered at the end of the term, as shown in table (5).

Table (5)
Correlation coefficients between aptitude results and achievement
tests scores

		Aptitude Test
Eng 120T	Pearson Correlation	.415
	Sig. (2-tailed)	.069
	N	20
Eng 118T	Pearson Correlation	.597**
	Sig. (2-tailed)	.005
	N	20
Eng 113T	Pearson Correlation	.368
	Sig. (2-tailed)	.110
	N	20
Eng 112T	Pearson Correlation	.100
	Sig. (2-tailed)	.676
	N	20
Eng 111T	Pearson Correlation	.444*
	Sig. (2-tailed)	.050
	N	20

** Correlation is significant at less than the 0.01 level.

* Correlation is significant at less than the 0.05 level (2-tailed)

As table (5) demonstrates, the majority of achievement tests correlate low with the aptitude test. The correlations range from very low to moderate low. The correlation is of .415 between 120 and aptitude, .368 between 113 and aptitude, .100 between 112 and aptitude and .444

between 111 and aptitude. The only significant and moderately high correlation is .59 between 118 and aptitude test. Therefore, at this early stage, the predictive validity of the aptitude test could not be confirmed as it was shown in the results of the present pilot study.

Chapter Four

Data Analysis

4.1 Introduction

This chapter presents the statistical analysis of the findings on the cloze test, aptitude test and the five achievement tests administered to the subjects. In this chapter, statistical procedures used in the data analysis are explained. The discussion is divided into four sections. First, reliability coefficients are computed for the study's variables. Second, the results of the descriptive statistics of the subjects' scores on the different variables are demonstrated. Third, inferential statistics are analyzed on the cloze test, aptitude test and achievement tests. Finally, regression coefficients for all variables are calculated. At the end of the chapter an overall discussion is provided.

4.2 Statistical Procedures Used

Several statistical procedures were used for the purpose of answering the questions posed earlier in this research. These statistics are as follows: first, several types of reliability values of the cloze and aptitude tests were computed. These reliability formulas are the Guttman, the Spearman-Brown prophecy formula, and the reliability coefficient of Cronbach Alpha. Such values measure tests' reliability

with respect to internal consistency reliability that could be determined from a single administration of a test.

The second statistical analysis utilized in this study showed the descriptive statistics such as means and standard deviations of the cloze, aptitude and the five achievement tests.

Then, inferential statistics were employed in order to make generalizations about the variables that were administered to a specific set of subjects. t- Test, ANOVA, and the Pearson product-moment correlation coefficient (r) were employed as procedures for inferential statistics. t- Test was utilized to compare the performance of the male subjects and female subjects on both the cloze and aptitude tests. ANOVA was employed to compare the performance of the subjects on the cloze test and on the aptitude test since the subjects were divided into more than two groups. This comparison was conducted to detect the differences in means among groups and within groups on the same test. Scheffe's test was used to pinpoint where the differences detected by the ANOVA occurred exactly. The Pearson's product-moment correlation was then applied to see how the two tests correlate with each other in order to be used as alternatives to predict future success. Then, it was employed to determine the interrelationships among the cloze test and the five achievement tests. As a final step, this

correlation was also conducted between the aptitude test and the same five achievement tests.

Finally, regression analysis was used to determine numerically the functional relationships between cloze test scores of the students, their aptitude test scores and their scores on their achievement tests.

4.3 Discussion of the Statistical Analysis

The following section displays the findings of the statistical analysis for the variables of the study. This analysis is divided into reliability coefficients, descriptive, inferential and regression analysis.

4.3.1 Several Reliability Coefficient Indices

For assessing the cloze and aptitude tests' reliability, the present study employed both the Guttman and the Spearman Prophecy Formulas, based on the split- half procedure. In addition, the reliability of Cronbach Alpha was calculated.

4.3.1.1. The Reliability of Cloze and Aptitude Tests (Cronbach Alpha)

Table (6) shows the reliability coefficients calculated using Cronbach Alpha of the cloze and aptitude tests.

Table (6)

The Reliability coefficients of cloze test and aptitude test (Cronbach Alpha).

Test	Reliability Coefficients		
	Male Students	Female Students	Both
Cloze Test	.867	.829	.862
Aptitude Test	.926	.898	.891
No. of Subjects	94	108	202

As can be seen from table (6), both the cloze and aptitude tests tend to produce high reliability estimates whether with the total group or with the female or male groups as separate groups. Therefore, it could be claimed that the cloze and aptitude tests are reliable tools to produce consistent results as measured by Cronbach Alpha.

4.3.1.2 The Reliability Coefficients of Cloze and Aptitude Tests

(Split- half technique)

Table (7)

The Reliability coefficients of cloze and aptitude tests using split-half procedure

Test	Reliability Coefficients					
	Guttman Split-half			Spearman-Brown Prophecy		
	Male Students	Female Students	Total Subjects	Male Students	Female Students	Total Subjects
Cloze T.	.774	.769	.783	.790	.776	.798
Apt. T.	.807	.705	.812	.843	.716	.829
No. of Subjects	94	108	202	94	108	202

Table (7) also gives reliability coefficients of the cloze and aptitude tests which were calculated by using split-half method adjusted by the Spearman-Brown formula and Guttman statistic. The Spearman –Brown Prophecy reliability coefficients of the cloze and aptitude tests indicate that the cloze and aptitude tests are reliable in the .798 to .829 range. The Guttman reliability coefficients of the two tests are similar to the Spearman-Brown Prophecy formula, though they are slightly lower ranging from .783 to .812. These values are acceptably high and are judged to be within acceptable limits even when considering the Guttman value that is known to show the lower bounds of reliability. However, the

aptitude test, with more items, yielded a higher reliability estimate than the cloze test. This is quite understandable since the reliability coefficient is affected by the number of items.

To conclude this section, it could be confirmed that the cloze and aptitude tests are reasonably reliable to produce consistent results. According to Kubiszyn and Borich (1987), the high reliability coefficient may be attributed to the clarity of test items and to the good way of writing these items, to the proper length of test items, or to the varied performance of the subjects. These researchers claim that heterogeneous subjects result in high reliability coefficients. For this reason, a display of descriptive characteristics, t- Test and Anova analyses were applied to see if a difference of performance exists between the groups.

4.3.2 Descriptive Statistics

The findings shown in table (8) are the descriptive statistics of the cloze test, aptitude test, and the five achievement tests including the number of subjects, means, standard deviations, minimum scores and maximum scores of both male and female students.

Table (8)
Means and standard deviations of cloze test, aptitude test and
achievement tests' scores

Test	Descriptive statistics				
	Mean	Standard Deviation	Minimum	Maximum	No.
Cloze Test	19.11	5.99	6	30	202
Aptitude Test	66.68	20.66	20	104	202
Achievement test Eng. 120	64.45	21.92	0	100	202
Achievement test Eng. 118	63.32	23.68	0	97	202
Achievement test Eng. 113	62.26	19.81	0	98	202
Achievement test Eng. 112	68.55	20.42	0	98	202
Achievement test Eng. 111	62.61	22.41	0	96	202

It is clear from the table above that the cloze mean score could be described as high with respect to other mean scores cited in previous studies. This relatively high mean score of the cloze test could be attributed to the careful selection of the passage in constructing the cloze test. As explained earlier, the readability level of the passage was in

accordance with students' reading ability. Therefore, the emphasis of several researchers on the careful selection and preparation of the cloze test is again confirmed by the present study's results. The mean score on the aptitude test is 66.68 (s. d. =20.66). As for the achievement tests, their mean scores are approximate. Eng 113 indicated the lowest mean score whereas Eng 112 calculated the highest mean score with a range from (62.26 to 68.55), all out of 100.

Table (9)

Means, and standard deviations of the cloze test, aptitude test and achievement tests' scores (Males)

Test	Descriptive statistics				
	Mean	Standard Deviation	Minimum	Maximum	No.
Cloze Test	17.19	6.35	6	30	94
Aptitude Test	53.95	18.28	20	95	94
Eng. 120	60.88	24.14	0	95	94
Eng. 118	58.84	28.13	0	95	94
Eng. 113	64.33	21.02	0	95	94
Eng. 112	64.30	23.53	0	98	94
Eng. 111	59.10	25.86	0	96	94

Table (10)

Means, and standard deviations of the cloze test, aptitude test and achievement tests' scores (Females)

Test	Descriptive statistics				
	Mean	Standard Deviation	Minimum	Maximum	No.
Cloze Test	20.79	5.14	9	30	102
Aptitude Test	77.86	15.58	44	104	102
Eng. 120	67.50	19.44	0	100	102
Eng. 118	67.22	18.25	15	97	102
Eng. 113	60.50	18.64	15	98	102
Eng. 112	72.42	16.26	0	98	102
Eng. 111	65.59	18.60	0	96	102

Tables (9) and (10) show the achievement level of female and male subjects as measured by three variables- the cloze, aptitude and achievement tests. Table (9) displays the descriptive statistics of male subjects whereas table (10) shows the descriptive characteristics of female subjects. This comparison was done in order to determine if a difference of performance existed between the two groups.

As shown in the tables, female subjects did better on the cloze and aptitude tests than the male subjects. Their mean score on the cloze test is 20.79 but the male subjects' mean score on the cloze test is 17.19. The difference between the means is 3.60 in favor of female subjects. The

mean score of female subjects on the aptitude test is 77.86, and the mean of male subjects on the same test is 53.95. The difference between the means on this variable is 23.91, also in favor of female subjects. Therefore, we can claim that on both tests female subjects' proficiency was higher than that of male subjects. Moreover, one can notice that the difference between the mean scores is greater on the aptitude test than on the cloze test.

The results of the achievement tests show a similar pattern of favor towards female students' proficiency except for the achievement test Eng. 113. On this achievement test male subjects' mean score is higher than female subjects' mean score. The difference between the mean in this case is 3.83. As a further step in the analysis and in order to check if the difference of means among the subjects is of statistical significance, t-Test was applied between female and male subjects as two independent groups.

4.3.3 Inferential Statistics

The following part of this chapter is divided into three main statistical analyses: t- Test, ANOVA, and the Pearson product-moment correlation coefficient r.

4.3.3.1 T- Test

Gosset, as cited in Spiegel (1991), worked on a refined technique known as t- Test in which the test is based on a null hypothesis. This null hypothesis says that sum of deviations of individual values from the mean value is equal to zero and thus there is no significant difference between groups. This study starts with this null hypothesis presuming that the difference between male scores and female scores on the cloze and aptitude tests is insignificant. To verify this point, t- Test was run on data obtained from the subjects as two independent groups, the male and female groups. Table (11), displays the values of t-Test statistical analysis and also the descriptive statistics of the two groups of the study, based on data obtained from the cloze test.

Table (11)
Male and female subjects' performance on the cloze test / T-test

Groups	N	Mean	Standard Deviation	T Value	Mean Difference	Significance
Male	94	17.19	6.35	-4.38	-3.60	.000
Female	108	20.79	5.14			

* P is significant at <0.01

** P is significant at <0.05

As can be seen from the table above, t value=-4.38 is statistically significant since level of significance (2-tailed) $=.000 < 0.05$. Therefore, t - Test reveals a statistically significant difference between female and male performance on the cloze test. This significant difference favors performance of female subjects over that of male subjects as it could be assumed from the mean difference.

The t - Test has also been applied in the following table to see if a significant difference of means exists between female and male subjects on the aptitude test.

Table (12)
Male and female subjects' performance on the aptitude test / t -Test

Groups	N	Mean	Standard Deviation	T Value	Mean Difference	Significance
Male	94	53.95	18.28	-10.015	-23.91	.000
Female	108	77.86	15.58			

* P is significant at < 0.01

** P is significant at < 0.05

From table (12) these results could be reported as follows: t value=-10.015 is statistically significant since level of significance (2-tailed) $= .000 < 0.05$, the level of confidence. t -Test also reveals that the mean of female subjects is significantly higher than that of male subjects.

To conclude the above section, on the basis of this analysis female and male subjects of first and second term of the academic year 1422-1423 H (2001-2002) should be considered as two separate groups and

the null hypothesis is rejected confirming a statistical significance between groups.

4.3.3.2 ANOVA

The subjects in this study were divided into four groups since a difference of English language proficiency among them was presumed to be significant. This division is based on the type of school attended (public or private) and on their extra training and exposure to English. Thus, the subjects divided could be listed as belonging to four groups as follows: 1) Group 1 includes students coming from public schools, 2) Group 2 includes students coming from private schools, 3) Group 3 includes students coming from public schools with extra training in English language not related to any school curriculum and 4) Group 4 to which belong students coming from private schools but with extra training in English other than the school curriculum. This division of subjects into four groups was applied both to male and female groups.

To investigate if the differences between the subjects' means on the cloze test and the aptitude test were significant, a one- way ANOVA test was conducted as shown in the tables below. The researcher started this statistical technique with the presumption that the difference between the samples and within the samples is significant. Therefore, Scheffe's test for post hoc comparisons was utilized for more research in the field of

subjects' performance. The following sub-sections will present these points.

4.3.3.2.1 ANOVA and Values of Scheffe F-test – Cloze test: male subjects

As mentioned earlier, determining if there are any significant differences between the means of male subjects, the Analysis of Variance (ANOVA) was used. The one-way ANOVA was calculated on the cloze test to measure these differences among the four groups of male subjects as shown in table (13).

The calculations of the ANOVA resulted in an F ratio of 11.139 indicating that there is a significant difference among the male subjects' performance on the cloze test at the 0.05 level of confidence. Similarly, Al- Fallay (1997a) found that there were statistically significant differences among the means of the subjects of different English proficiency levels in his research that investigated the reliability and validity of the fixed ratio multiple choice cloze test.

Table (13)

ANOVA (One-way Analysis of Variance)

Cloze Test: male subjects

	SS	DF	MS	F	Sig.
Between groups	1013.906	3	337.969	11.139	.000
Within Groups	2730.647	90	30.341		
Total	3744.553	93			

Then, the Scheffe's test for post hoc comparisons was computed in the following table to show, in specific, the groups who were responsible for the differences among the male subjects' performance on the cloze test.

Table (14)

Values of Scheffe F-test of the difference between the male subjects' means on the cloze test

Comparisons		Mean difference	Sig.
1	2	-4.30	.093
	3	-7.39*	.001
	4	-11.51*	.002
2	1	4.30	.093
	3	-3.10	.599
	4	-7.21	.163
3	1	7.39*	.001
	2	3.10	.599
	4	-4.11	.653
4	1	11.51*	.002
	2	7.21	.163
	3	4.11	.653

* The mean difference is significant at the .05 level.

Scheffe's test revealed many results as follows: first, it showed that the difference between Group 1 (male students who studied in public schools with no extra training in English language) and Group 2 (male students coming from private schools) is statistically insignificant, with a mean difference of -4.30. Second, the difference between Group 1 (public school male students) and Group 3 (male students from public schools

with extra training in English) has proved to be statistically significant, with a mean difference of -7.39 in favor of Group 3. Third, the analysis detected also a statistically significant difference between Group 1 and Group 4 (male private school students but with additional training in English), with a mean difference of -11.51 in favor of the last group. Fourth, the difference between Group 2 and Group 1, 3 and 4 was computed to be statistically insignificant. Fifth, the analysis could not indicate statistically significant difference between Group 2 and 3 and between Group 2 and 4. More analysis was given for computing the difference between Group 3 and all other Groups. In this case the statistically significant difference was only found between Groups 3 and 1, with a mean difference of 7.39 in favor of Group 3. Finally, as for Group 4, the only statistically significant difference was found between that Group and Group 1, with a mean difference of 11.51, whereas the mean differences between Group 4 and all the other two Groups were statistically insignificant.

4.3.3.2.2 ANOVA and Values of Scheffe - test – Aptitude Test: male students

Since it was assumed that the acquisition of a second language may have an impact on first language analysis and since the groups varied in their level of English proficiency, it was important to find out if the

aptitude test would discriminate equally well among them. Thus, as a further step, ANOVA was computed to determine the significance of the difference among means of the male subjects on the aptitude test.

Table 15 displays the results of one-way ANOVA on the aptitude test. It is clear from the table below that there is a significant difference among the groups' performance on the aptitude test at the 0.05 level of confidence, since the F-ratio is 4.299.

Table (15)

ANOVA (One-way Analysis of Variance)

Aptitude test: male students

	SS	DF	MS	F	Sig.
Between groups	3893.223	3	1297.741	4.299	.007
Within Groups	27167.511	90	301.861		
Total	31060.734	93			

Post hoc comparisons were applied using the Scheffe's test for the second time to pinpoint which groups are the sources of the difference in the male subjects' performance on the aptitude test. The following table presents the findings of this statistical method.

Table (16)

Values of Scheffe F-test of the difference between the male subjects' means on the aptitude test

Comparisons		Mean difference	Sig.
1	2	-5.92	.739
	3	-11.38	.264
	4	-27.63*	.028
2	1	5.92	.739
	3	-5.46	.899
	4	-21.71	.197
3	1	11.38	.264
	2	5.46	.899
	4	-16.25	.467
4	1	27.63*	.028
	2	21.71	.197
	3	16.25	.467

* The mean difference is significant at the .05 level

The Scheffe analysis applied here reveals many results: First, it shows that the difference between Group 1 and Group 2 is statistically insignificant. Second, the difference between Group 1 and Group 3 is also statistically insignificant. However, the difference of means between Group 1 and Group 4 is statistically significant, with a mean difference of -27.63 in favor of the last Group. Third, the difference of means between Group 2 and 1, 3 and 4 is statistically insignificant. As for Group 3, the difference of means between this Group and 1, 2, 4 Groups is statistically

insignificant. Finally, regarding Group 4, the only statistically significant difference of means exists between Group 4 and Group 1 whereas the difference between Group 4 and Groups 2 and 3 is statistically insignificant. Therefore, it could be claimed here that the Scheffe analysis could detect only one statistically significant difference existing between Group 1 and Group 4 and vice versa, but it failed to detect significant differences among Groups 1, 2 and 3 and vice versa, using their scores on the aptitude test.

4.3.3.2.3 ANOVA and Values of Scheffe F-test- Cloze test: female subjects

The same statistical procedure was employed for the purpose of detecting differences of means between groups of female subjects. The following tables present these points.

Table (17)

ANOVA (One-way Analysis of Variance)

Cloze Test: female subjects

	SS	DF	MS	F	Sig.
Between groups	279.195	3	93.065	3.803	.012
Within Groups	2544.907	104	24.470		
Total	2824.102	107			

This procedure resulted in an F ratio of 3.803 indicating that there is a significant difference at the 0.05, level of confidence among the four groups in their performance on the cloze test. Therefore, Scheffe's post hoc comparisons were then conducted in table (18).

Table (18)

Values of Scheffe F-test of the difference between the female subjects' means on the cloze test

(I) Study	(J) Study	Mean difference	Sig.
1	2	-2.88	.127
	3	-1.48	.777
	4	-4.52*	.047
2	1	2.88	.127
	3	1.41	.852
	4	-1.64	.828
3	1	1.48	.777
	2	-1.41	.852
	4	-3.04	.462
4	1	4.52*	.047
	2	1.64	.828
	3	3.04	.462

* The mean difference is significant at the .05 level.

The Scheffe' analysis reveals the following results: first, it indicates that the difference of means between Group 1 and Groups 2 and 3 is statistically insignificant, whereas the difference between Group 1 and

Group 4 is statistically significant. Second, the difference between Group 2 and all other Groups is statistically insignificant. Third, the difference between Group 3 and Groups 1, 2 and 4 is also statistically insignificant. Finally, the difference between Group 4 and Group 1 is statistically significant, with a mean difference of 4.52 in favor of Group 4, whereas the difference between Group 4 and Groups 2 and 3 is statistically insignificant.

As it was performed earlier with male subjects, a one-way ANOVA was also run on the means of female subjects' aptitude tests. The results are displayed in table (19).

Table (19)

ANOVA (One-way Analysis of Variance)

Aptitude test: female students

	SS	DF	MS	F	Sig.
Between groups	189.465	3	63.155	.255	.858
Within Groups	25525.433	103	247.820		
Total	25714.897	106			

This analysis resulted in an F ratio of .255 indicating that there is no statistically significant difference among groups of female subjects.

Therefore, it could be stated here that ANOVA failed to detect variations among female groups, based on their scores on the aptitude test.

4.3.3.3 Pearson Product-Moment Correlation Coefficient r

In this section, the present researcher is interested in showing the extent to which subjects' position in one distribution, i. e. the cloze test, is similar or dissimilar to his or her position in a different distribution, i. e. the aptitude test and achievement tests. Therefore, Pearson's product-moment correlation (r) was performed to indicate the degree of commonality among the tests used in the present study. Such correlation helps telling, at a glance, the strength and direction of the relationship between distributions. Kubiszyn and Borich point out that "correlation coefficient ranges from -1.0 to +1.0 and the closer a coefficient gets to -1.0 or to +1.0, the stronger the relationship. The sign of the coefficient tells us whether the relationship is positive or negative" (Kubiszyn and Borich, 1987: 264). Based on these statistical principles, the correlation coefficients between the cloze and aptitude tests, between the cloze and achievement tests and between the aptitude and achievement tests were calculated as shown in the subsections below.

4.3.3.3.1 Correlation between Cloze and Aptitude Tests

Correlation coefficient between cloze and aptitude tests was calculated in order to determine the degree of interrelationship between them. This correlation was run to include the total subjects' scores. Table (20) displays it.

Table (20)

Correlation coefficient between cloze test results and aptitude test scores

		Cloze T.
Apt. T.	Pearson Correlation	.650*
	Sig. (2-tailed)	.000
	N	202

*Correlation is significant at less than the 0.05 level (2-tailed)

**Correlation is significant at less than the 0.01 level (2-tailed)

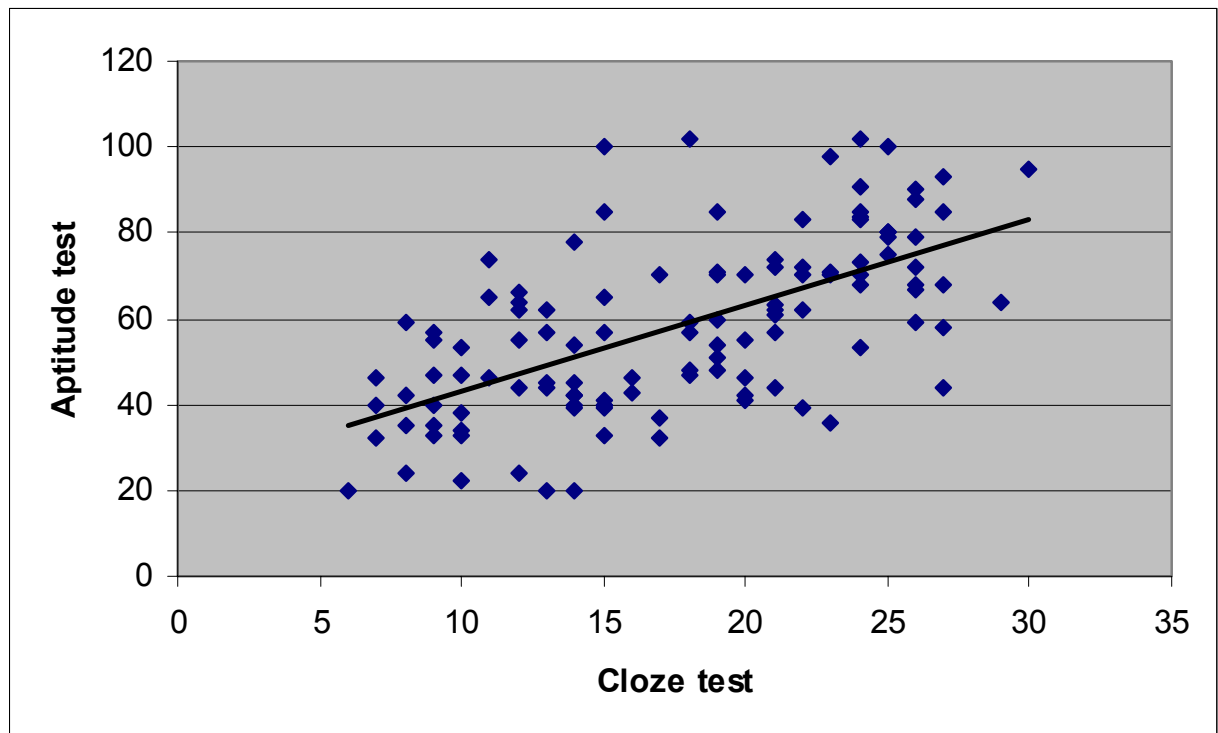
According to table (20), a significant correlation exists between cloze and aptitude tests, indicating a commonality between the two tests. The value for the correlation between the cloze and aptitude tests is $r=.65$, significant at $P<.01$.

The following Scatter Diagram is obtained by plotting the points that show the total students' scores on the cloze test and their scores on the aptitude test. It shows that the data of the study appear to be approximated by a straight line. Thus, a linear relationship exists between the variables. In such a case, the students' scores on the aptitude test tend

to increase as the scores on the cloze test increases. This correlation is considered as positive or direct correlation.

Figure 1

Correlation between the cloze test scores and the aptitude test scores



4.3.3.3.2 Correlation Coefficients between Cloze and Achievement

Tests: male subjects

This part checks the validity of the cloze test to be used as a predictor of future achievement test scores. Therefore, the Pearson product moment correlation was conducted in order to see the degree of interrelationships that holds between cloze and achievement tests, based on male subjects' scores. The divisions of male subjects into four groups

have not been taken into account for the purpose of doing such correlation since the only significant difference existed between Group 1 and Groups 3 and 4. Moreover, the number of students belonging to these groups was very minor to perform correlations. Therefore, the following tables demonstrate the results of the total number of male students.

Table (21)
Correlation coefficients between scores of cloze test and scores of
achievement tests of male subjects

		Cloze T.
Eng.120T	Pearson Correlation	.457**
Vocabulary	Sig. (2-tailed)	.000
Building	N	94
Eng.118T	Pearson Correlation	.408**
Translation	Sig. (2-tailed)	.000
	N	94
Eng.113T	Pearson Correlation	.519**
Reading	Sig. (2-tailed)	.000
Comprehension	N	94
Eng.112T	Pearson Correlation	.638**
Listening &	Sig. (2-tailed)	.000
Speaking	N	94
Eng.111T	Pearson Correlation	.387**
Basic	Sig. (2-tailed)	.000
Language Skills	N	94
Total Ach.	Pearson Correlation	.547**
	Sig. (2-tailed)	.000
	N	94

*Correlation is significant at less than level 0.05

** Correlation is significant at less than level 0.01

As shown in table (21), the correlation coefficient of the total achievement tests with the cloze test is 0.547. This reflects a positive and

moderately high correlation. Similarly, performance on the cloze test correlates positively and significantly with the different achievement tests. All the correlation coefficients are significant at less than the 0.01 level. Each achievement test correlates with the cloze test as follows: achievement test Eng.120: ($r=0.457$), achievement test Eng.118: ($r=0.408$), achievement test Eng.113: ($r=0.519$), achievement test Eng.112: ($r=0.638$) and achievement test Eng.111: ($r=0.387$). So, on the basis of these statistical analyses, it could be discovered that there is a positive correlation between the male subjects' scores on the cloze test and their scores on their achievement tests administered at the end of term 1 of university study. In addition, one can notice that the highest correlation coefficient was found between the cloze test and Eng.112 that teaches listening and speaking, whereas the lowest correlation was between cloze test and Eng.111, the course of basic language skills.

4.3.3.3 Correlation Coefficients between Aptitude and Achievement Tests: male subjects

As for the aptitude test, similar correlations to those performed in the previous section, were calculated between male subjects' scores on the aptitude test and their scores on their achievement tests. The following table displays the results.

Table (22)
Correlation coefficients between aptitude results and achievement tests scores of male subjects

		Apt. T.
Eng.120T	Pearson Correlation	.520**
	Sig. (2-tailed)	.000
	N	94
Eng.118T	Pearson Correlation	.501**
	Sig. (2-tailed)	.000
	N	94
Eng.113T	Pearson Correlation	.555**
	Sig. (2-tailed)	.000
	N	94
Eng.112T	Pearson Correlation	.663**
	Sig. (2-tailed)	.000
	N	94
Eng.111T	Pearson Correlation	.447**
	Sig. (2-tailed)	.000
	N	94
Total Ach.	Pearson Correlation	.617**
	Sig. (2-tailed)	.000
	N	94

* Correlation is significant at less than level 0.05

** Correlation is significant at less than level 0.01

Male students' scores on the aptitude test correlate positively with the set of scores of the achievement test. The coefficient is high and

significant at less than the 0.01 level. The details of how each achievement test correlates with the aptitude test is explained as follows: achievement test Eng.120: ($r=.520$), achievement test Eng.118: ($r=.501$), achievement test Eng.113: ($r=.555$), achievement test Eng.112: ($r=.663$), and achievement test Eng.111: ($r=.447$). These results confirm the validity of the aptitude test as a tool to predict male students' future success. However, the highest correlation was between aptitude and Eng. 112, the course of listening and speaking whereas the lowest was found between aptitude and Eng.111, the course of basic language skills. One can note here that the correlations obtained between the aptitude and achievement tests are higher than those obtained between the scores on cloze and achievement tests. The correlations between the cloze and achievement tests range from .387 to .638, whereas the correlations between the aptitude test and the five achievement tests range from .447 to .663.

4.3.3.3.4 Correlation Coefficients between Cloze and Achievement Tests: female subjects

The same analyses were run on the sets of scores of female students. The division of female subjects into four groups was neglected when doing correlations since only one statistically significant difference has proved to be between all Groups 1,2,3 and Group 4. In addition, the number of Group 4 students was too small to perform such correlational

analysis. So, the results shown in the following tables include all groups of female students.

Table (23)

Correlation coefficients between scores of cloze test and scores of achievement tests of female subjects

		Cloze T.
Eng.120T	Pearson Correlation	.602**
	Sig. (2-tailed)	.000
	N	108
Eng.118T	Pearson Correlation	.649**
	Sig. (2-tailed)	.000
	N	108
Eng.113T	Pearson Correlation	.714**
	Sig. (2-tailed)	.000
	N	108
Eng.112T	Pearson Correlation	.550**
	Sig. (2-tailed)	.000
	N	108
Eng.111T	Pearson Correlation	.435**
	Sig. (2-tailed)	.000
	N	108
Total Ach.	Pearson Correlation	.696**
	Sig. (2-tailed)	.000
	N	108

*Correlation is significant at less than level 0.05

** Correlation is significant at less than level 0.01

As shown in table (23), performance of female students on the cloze test correlates positively and significantly with their total performance on the end of term achievement tests. The coefficient is high and significant at less than the 0.01 level. Each achievement test correlates with the cloze test as follows: Eng.120: $r=.602$; Eng.118: $r=.649$; Eng.113: $r=.714$; Eng.112: $r=.550$ and Eng.111: $r=.435$. Notice that the correlation coefficients between cloze and achievement tests, based on female students' performance, are higher than those of their peers at the male section of the English Department. Consequently, it could be claimed that the cloze test is a better predictor of future success of female students than it is for male students. The highest correlation is clearly shown between cloze and Eng.113, the course of reading comprehension, whereas the lowest correlation exists between cloze and Eng.111, the course of basic language skills.

4.3.3.3.5 Correlation Coefficients between Aptitude and Achievement

Tests: female subjects

Correlations similar to those done in the previous sections were calculated between female students' scores on the aptitude and achievement tests. The results are shown in table (24).

Table (24)

Correlation coefficients between scores of aptitude test and scores of achievement tests of female subjects

		Apt. T.
Eng.120T	Pearson Correlation	.297**
	Sig. (2-tailed)	.002
	N	108
Eng.118T	Pearson Correlation	.389**
	Sig. (2-tailed)	.000
	N	108
Eng.113T	Pearson Correlation	.351**
	Sig. (2-tailed)	.000
	N	108
Eng.112T	Pearson Correlation	.191
	Sig. (2-tailed)	.056
	N	108
Eng.111T	Pearson Correlation	.387**
	Sig. (2-tailed)	.000
	N	108
Total Ach.	Pearson Correlation	.390**
	Sig. (2-tailed)	.000
	N	108

*Correlation is significant at less than level 0.05

** Correlation is significant at less than level 0.01

As can be noticed in table (24), the correlation coefficient of the total achievement test scores with the aptitude test is 0.390. This reflects a

positive, but not very high correlation. Moreover, the results reveal that only four achievement tests reflect a significant correlation. These correlations reported in the above table are low, even though they are statistically significant at $P < .05$ level of significance. A detailed account of how each achievement test correlates with the aptitude test based on female students' scores could be provided as follows: Eng.120: $r = .297$; Eng.118: $r = .389$; Eng.113: $r = .351$ and Eng.111: $r = .387$, but Eng.112 does not show a significant correlation. These correlations between aptitude and achievement tests of female students are lower than those of the male group. The correlations between aptitude and achievement tests of male students range from .447 to .663, whereas those of female subjects range from .297 to .389. Consequently, it could be proposed here that the aptitude test is a better predictor of male subjects' scores than it is of female students' achievement. Moreover, the aptitude does not seem to follow a predictable pattern of female achievement of Eng.112, the course of listening and speaking.

4.4 Regression

As proven in the section above, a positive and significant correlation exists between cloze and achievement tests, and between the majority of achievement tests and the aptitude test, it was appropriate to perform regression analysis on these variables. A researcher such as

Spiegel, explained that "regression analysis permits to estimate the value of a variable Y corresponding to a given value of a variable X", (Spiegel, 1991: 268). According to him, X and Y could represent dependent or independent variables, based on the purpose of the study. For the present study, the cloze and aptitude tests are considered to be the independent variables whereas achievement tests are the dependent variables. In regression equation the dependent variable- the achievement tests are estimated from the independent variables- the cloze and aptitude tests. The cloze, aptitude and achievement tests are measured from their means, and Beta is the parent value of the regression.

This part of analysis is divided into two sections. The first considers regression analyses based on male students' scores while the second is devoted to female subjects' scores.

4.4.1 Regression analysis: male students

The following tables show how each achievement test score could be predicted from cloze and aptitude tests scores, based on performance of the male group.

Table (25)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.538 a	.290	.274	20.57

a. Predictors: (constant), Cloze T., Aptitude T.

Table (26)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	20.723	7.056		2.937	.004
Cloze T.	.722	.469	.190	1.539	.127
Apt. T.	.509	.161	.390	3.168	.002

Dependent Variable: Eng.120 Total

The above tables illustrate how well the achievement test of E120 could be predicted from cloze and aptitude tests' scores. As it can be seen from tables (25) and (26), the relationship is significant only with the aptitude test scores, accounting for 29% of the variance in the E120 scores. Moreover, an equation could be provided permitting the prediction of E120 scores based on the aptitude test as follows:

$$\text{Eng.120} = 20.72 + .509 \text{ Apt. T.}$$

In order to discover how Eng.118 scores can be predicted by cloze and aptitude tests' scores, a regression analysis was performed with the cloze and aptitude tests, as independent variables and Eng.118 as the dependent variable regarding male group performance. The tables below represent an account of this analysis.

Table (27)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.509 a	.259	.243	24.48

a. Predictors: (constant), Cloze T., Aptitude T.

Table (28)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	14.897	8.261		1.803	.075
Cloze T.	.538	.549	.121	.980	.330
Apt. T.	.643	.191	.418	3.370	.001

Dependent Variable: Eng.118 Total

Similarly, table (27) and (28) report a statistically significant value only for the aptitude test, accounting for 25% of the variance in Eng.118 scores. This equation could be provided permitting the prediction of Eng.118 scores from aptitude test scores: $\text{Eng.118} = 14.89 + .643 \text{ Apt. T.}$

The following tables show the regression coefficient of the cloze test and the aptitude test scores with the achievement test scores of the English course 113- Reading Comprehension.

Table (29)
Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.587 a	.344	.329	17.21

a. Predictors: (constant), Cloze T., Aptitude T.

Table (30)
Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	26.324	5.843		4.505	.000
Cloze T.	.863	.390	.261	2.212	.030
Apt. T.	.429	.134	.376	3.193	.002

a. Dependent Variable: Eng.113 Total

Table (29) and (30) reflect a statistically significant relationship that holds among cloze, aptitude and Eng.113 tests scores, accounting for 34% of the variance in Eng.113 scores. Therefore, on this basis, Eng.113 total scores could be predicted as follows:

Eng.113 Total= 26.32+.86 Cloze T.

Eng.113 Total= 26.32+ .429 Apt. T.

The same analysis has also been computed between cloze, aptitude and achievement test Eng.112 to demonstrate how Eng.112 total scores could be predicted from cloze and aptitude test scores of male students.

The following tables show the calculation.

Table (31)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.709 a	.503	.492	16.77

a. Predictors: (constant), Cloze T., Aptitude T.

Table (32)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	12.946	5.663		2.286	.025
Cloze T.	1.283	.377	.347	3.401	.001
Apt. T.	.545	.131	.425	4.169	.000

a. Dependent Variable: Eng.112 Total

The R Square of .503 reported in tables (31) and (32), reveals a statistical significance, accounting for 50% of the variance in the Eng.112 total scores, as well as levels of significance of cloze and aptitude tests scores. Two equations could be given as follows:

Eng.112 Total= 12.94+1.28 Cloze T.

Eng.112 Total=12.94+.545 Apt. T.

Finally, Eng.111 scores as a dependent variable, cloze test and aptitude test scores as independent variables were analyzed to find the regression coefficients among them. The following tables display this procedure.

Table (33)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.461 a	.213	.195	23.21

a. Predictors: (constant), Cloze T., Aptitude T.

Table (34)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	21.962	8.072		2.721	.008
Cloze T.	.634	.532	.154	1.190	.237
Apt. T.	.479	.181	.342	2.640	.010

Dependent Variable: Eng.111 Total

As shown above, R Square=.213, accounting for 21% of the variance in Eng.111 total scores. The relationship is significant only between aptitude and Eng.111 tests' scores. The cloze test scores are unable to predict Eng.111 Total scores, since there is no statistical

significance between them. On this basis, Eng.111 total scores could be predicted as follows: $\text{Eng.111 Total} = 21.96 + .479 \text{ Apt. T.}$

4.4.2 Regression Analysis: Female Students

The same analysis was run on the different variables, but this time female students' scores were taken into consideration. The following tables display how well Eng.120 total score could be predicted from cloze and aptitude tests' scores of female subjects.

Table (35)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.604 a	.365	.353	15.70

a. Predictors: (constant), Cloze T., Aptitude T.

Table (36)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	20.657	8.269		2.498	.014
Cloze T.	2.327	.347	.613	6.702	.000
Apt. T.	-2.296E-02	.115	-.018	-.200	.842

Dependent Variable: Eng.120 Total

As shown above, R Square=.365, accounting for 36% of the variance in Eng.120 total scores, the relationship is significant. However, level of significance was only found between cloze and Eng.120 tests' scores. The aptitude test scores are unable to predict Eng.120 Total scores, since there is no statistical significance between them. On this basis, Eng.120 total scores could be predicted as follows:

$$\text{Eng.120 Total} = 20.65 + 2.32 \text{ Cloze T.}$$

A second analysis is displayed in the following tables to highlight coefficients among the same independent variables and Eng.118 total scores of female students.

Table (37)
Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.651 a	.424	.413	14.03

a. Predictors: (constant), Cloze T., Aptitude T.

Table (38)
Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	15.695	7.311		2.147	.034
Cloze T.	2.172	.309	.612	7.023	.000
Apt. T.	8.249E-02	.102	.070	.805	.423

Dependent Variable: Eng.118 Total

As for Eng.118 total scores, the relationship is only significant, accounting for 42% of the variance in the achievement test of Eng.118, with the cloze test. Level of significance does not allow the prediction of Eng.118 total scores from aptitude test scores. So, the equation can be given as follows:

$$\text{Eng.118} = 15.69 + 2.17 \text{ Cloze T.}$$

The following tables show the regression coefficient of the cloze test and the aptitude test scores with the achievement test scores of the English course 113- Reading Comprehension, of female students.

Table (39)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.714 a	.510	.501	13.22

a. Predictors: (constant), Cloze T., Aptitude T.

Table (40)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	8.284	6.889		1.203	.232
Cloze T.	2.643	.291	.729	9.068	.000
Apt. T	-3.442E-02	.097	-.029	-.356	.722

a. Dependent Variable: Eng.113 Total

As displayed in tables (39) and (40), the relationship is significant, accounting for 51% of the variance in Eng.113 total scores only with cloze test scores. The equation can be written as follows:

$$\text{Eng.113} = 8.28 + 2.64 \text{ Cloze T.}$$

The same analysis has also been computed between cloze, aptitude and achievement test Eng.112 to demonstrate how Eng.112 total scores could be predicted from cloze and aptitude test scores. The following tables show the calculation.

Table (41)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.559 a	.312	.298	13.69

a. Predictors: (constant), Cloze T., Aptitude T.

Table (42)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	41.311	7.413		5.572	.000
Cloze T.	1.940	.309	.607	6.269	.000
Apt. T.	-.119	.101	-.114	-1.175	.243

Dependent Variable: Eng.112 Total

The results reported in the above tables show that the relationship is significant, accounting for 31% of the variance in Eng.112 scores. Moreover, it illustrates how closely the cloze test scores can predict the term grade of Eng.112 based on the following equation:

$$\text{Eng.112} = 41.31 + 1.94 \text{ Cloze T.}$$

Finally, Eng.111 scores as a dependent variable, cloze test and aptitude test scores as independent variables were analyzed to find the regression coefficients among them. The following tables display this result.

Table (43)

Regression – Model Summary

Model	R	R square	Adjusted R square	Standard Error of the Estimates
1	.476 a	.226	.211	16.59

a. Predictors: (constant), Cloze T., Aptitude T.

Table (44)

Regression Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
(Constant)	20.341	8.792		2.313	.023
Cloze T.	1.168	.367	.323	3.182	.002
Apt. T.	.266	.122	.222	2.188	.031

Dependent Variable: Eng.111 Total

The results shown in tables (43) and (44), reveal that the relationship is significant with both cloze and aptitude tests, accounting for 22% of the variance in Eng.111 total scores. Two equations could be given as follows:

$$\text{Eng.111} = 20.34 + 1.168 \text{ Cloze T.}$$

$$\text{Eng.111} = 20.34 + .266 \text{ Apt. T.}$$

4.5 Discussion of the Results

This section discusses the results of the data analyses in light of some possible explanations coming from the existing literature mentioned earlier. The first and very important issue that could be considered here relates to the cloze mean score of the present study that is not in line with other cloze mean scores shown in previous studies (Hanania and Shikhani, 1986; Fotos, 1991). Fotos (1991), concluded that "cloze test mean scores are rather low regardless of the institution, the language tested, or the type of students, especially when exact scoring is used" (Fotos, 1991: 329). The high cloze mean score of the present study of both male and female students could be explained in terms of several factors. One of these factors relates to the careful selection of the possible passage that suited the readability level of the participants of this study. Therefore, the results of this study are in accordance with Brown recommendations for the selection of the passage to be used for the cloze

procedure. For him, "tester/ researcher intervention is necessary for developing sound cloze test" (Brown, 1993: 110).

Another factor that would explain the high cloze mean score could be the form of the cloze test used here. The present study used a matching cloze test that provides students with a list of words from which they fill in the blanks of the passage. This form of the cloze test is less difficult than the open-ended cloze test. According to Baldauf and Propst, there are several reasons to use this form with beginning level EFL students. For example, "beginning level ESL students need to use more text information when they are reading than do students who are more proficient in English. This additional information is given in the form of word choices for students to insert into the blanks in the test they are reading" (Baldauf and Propst, 1979: 5). Moreover, "In matching cloze, unlike some multiple choice cloze tests, the directions require students to actually write a word into each of the blanks in the text so that they can later use that as text information in determining what words go in the other blanks" (Baldauf and Propst, 1979: 6). In addition, Goodman as cited in Baldauf and Propst (1979), also pointed out that matching cloze tests engage beginning level of EFL students in an "editing" process as they read. This editing process works on the principle of the hypothesis the reader employs to fill in the blanks of the passage. As long as the words provided can be used only once, the reader most of the time is

revising his or her hypotheses about the correct word to select to fill in the blanks. This process of reading and rereading and forming hypotheses reflects considerable evidence that the testee is editing what is being read and that he or she is employing several reading strategies and linguistic competence to fill in the words of the matching cloze test. Thus, the choice of this form to use in this study has shown its significance at both the theoretical and practical levels.

A final point that might also explain the high mean score of the matching cloze test, in contrast with the low mean score of the open-ended cloze test, could be based on Al-Fallay's (1997) conclusion. According to him, "the cloze test and the M-C cloze test and the fact that the latter is a modified form of the former should not mislead the researcher into generalizing findings of cloze test's research to situations where the M-C cloze test is used" (Al-Fallay, 1997: 520). So, for the purpose of the present study, it would be advisable to regard the open-ended cloze test and the matching cloze test as two autonomous tests belonging to the same procedure, and each test needs investigations on its own.

Besides these explanations and arguments for using the matching cloze test, the procedure has proved to be reliable. Considering the number of items contained in the matching cloze test, it could be claimed that this form is easier to administer and to score than the aptitude test

used as the criterion measure in this study. This matching cloze test took only 35 minutes to administer and about one or two minutes to score per test, whereas the aptitude test took about 70 minutes to administer and about 15 minutes to score per test. Surprisingly, it should be emphasized here that with only 30 items, the matching cloze test maintains a very respectable degree of internal reliability, as can be seen in tables (6) and (7). Notice how the range between reliability values of cloze and aptitude tests is very narrow although the aptitude test has 120 items. Moreover, it is interesting to note that the reliability indices of the cloze and aptitude tests based on male scores are higher than those based on female students' scores. According to many scholars such as Kubiszyn and Borich (1987), these higher reliability values could be attributed to the group variability of male students who were more heterogeneous in their proficiency than female students. Anova results showed a statistically significant difference of means between Groups 1, 3 and 4 of male students, based on cloze test scores, and it showed a statistically significant difference of means between Groups 1 and 4 based on aptitude test scores. As for the female group, it could be claimed that despite the restricted range of the cloze and aptitude tests scores found in the homogenous group of female subjects, the fairly high reliability estimates reported indicate that the cloze and aptitude tests are reasonably reliable.

With respect to the cloze and aptitude tests used in this study as predictive tools of EFL students' success, a detailed explanation could be given as follows. First, as regards the cloze and aptitude tests and for the purpose of determining how strong the interrelationship that holds between the two tests, the whole sample size of the study was used. This exception is based upon the claims of many researchers such as Guilford who states that "the larger the sample size, the lower the correlation required to claim significance" (Guilford, 1965: 580 as quoted in Al-Khalaf (2000) research). Results show that the matching cloze test correlates positively and significantly with the aptitude test. The correlation coefficient between scores on the matching cloze and scores on the aptitude is positive, significant and relatively strong ($r=0.650$; $P<0.01$). Therefore, considering these results it could be suggested that the matching cloze test could be used as an alternative to the aptitude test to predict EFL students' future success. Thus, time is shortened and effort is lessened to make valid predictions of EFL students' success by using the matching cloze test.

Second, as for the relationship between cloze and achievement tests, several remarks could be provided. Correlations between the matching cloze and achievement tests are all statistically significant at $P<0.05$ level of significance or better. The matching cloze test scores and the total achievement tests scores shared a correlation coefficient of 0.696

based on female students' scores and a correlation coefficient of 0.547 considering male students' scores. More specifically, the correlation coefficients between cloze and achievement tests based on female students' scores ranged from relatively strong ($r=0.714$; $P<0.01$) to relatively moderate ($r=0.435$; $p<0.01$). But on the basis of male students' scores, they ranged from moderate ($r=0.638$; $P<0.01$) to relatively weak ($r=0.387$; $P<0.01$). From these correlation coefficients, it is obvious that the matching cloze test seems to be a better predictor of EFL future success of female students than it is of male students. This point could be partially explained by the fact that for a group with more experience in EFL, the matching cloze test constructed from EFL materials is supposed to be better predictor than it would be for a group with the less experience in EFL. A closer look at the educational background of female students reveals that female students had gained more experience in EFL than their counterparts at the male section. For example, 38.8% of the female students (42 out of 108) studied at both private and public schools whereas only 19.1% of the male subjects (18 out of 94) studied at private and public schools. Therefore, we can easily imagine that the female students were the group with the more exposure to EFL than the male students.

A final remark here should be made concerning the correlations between the aptitude and the end of term achievement tests. Correlations

between the aptitude and the total achievement tests scores are all statistically significant at $P < 0.01$ level of significance. The aptitude test correlated at 0.390 with the total achievement tests scores with the female students, and it correlated at 0.617 with the same variables based on male students' scores. In order to make a full comparative analysis of correlation coefficients between the aptitude and the end of semester achievement tests, it would be seem appropriate to consider the range of correlations between these two variables with both male and female performance. The correlation coefficients between aptitude and the end of term achievement tests ranged from very weak and insignificant ($r = 0.191$; $P > 0.05$) to relatively moderate ($r = 0.389$; $P < 0.01$) based on female students' scores, and they ranged from relatively strong ($r = 0.663$; $P < 0.01$) to moderate ($r = 0.447$; $P < 0.01$). In light of the above results, it is clear that the aptitude test seems to be a better predictor of male students' success than it is of female subjects. This conclusion confirms the above mentioned point stating that with the group with less previous EFL experience, the aptitude test constructed in their native language is a better predictor than any other test written in a foreign language such as English. These results are consistent with Al-Khalaf (2000)'s conclusions which showed that the aptitude test scores yielded the highest correlation coefficients with the end of term achievement tests for first year students at both the College of Education and the College of Art. Since students

from the first year presented the group with the least experience and training in EFL, the aptitude test predicted best these subjects' achievements.

Chapter Five

Conclusions

5.1 Introduction

As noted in the previous chapter, there is a statistical significance of cloze and aptitude tests as predictors of students' success and achievement in EFL courses. This chapter will discuss the study's conclusions based on the statistical results of the variables, implications of the study and suggestions for further investigation.

5.2 Conclusions

The purpose of this study was to investigate the possibility of using cloze test as a predictive tool of students' achievement and success in EFL courses. The first null hypothesis states that there is no correlation between the students' scores on the cloze test and scores of the same students on the end of semester achievement tests. Results show that scores on the cloze test used here correlate positively with scores of male and female students on their total end of semester achievement tests. All correlations are significant but they were higher with female students than with male students, as can be seen in tables (21) and (23). Therefore, the first null hypothesis is rejected.

The second null hypothesis states that there is no correlation between the students' scores on the aptitude test and scores of the same

students on the total end of semester achievement tests. Results show that the correlation that exists between aptitude and the total achievement tests are positive and significant, based on female and male students' scores. All correlations are significant but they were higher with male students than with female students, as can be seen in tables (22) and (24). Therefore, the second null hypothesis is rejected.

The third null hypothesis states that there is no correlation between cloze and aptitude tests' scores. Results show that the matching cloze test used here correlates positively and significantly with the aptitude test, based on the scores of the total number of subjects who participated in the study. The correlation coefficient between scores on the cloze and scores on the aptitude is positive, significant and relatively high ($r = 0.650$; $P < 0.01$) although the sample size is large. Therefore, the third null hypothesis is rejected, and it could be claimed with confidence that these two measures reveal an interesting pattern of interrelationships, as can be seen in table (20).

As for the five null sub hypotheses, a detailed explanation could be provided as follows: first, the first null sub hypothesis made earlier states that there is no correlation between the students' scores on the cloze and achievement test of Eng.120, and between the same students' scores on the aptitude and achievement test of Eng.120. The results show that for both male and female groups, scores on the cloze correlate positively and

significantly with scores on the achievement test of Eng.120. The correlation coefficients are .457 and .602 for the male and female students, respectively, as can be seen in tables (21) and (23). Results in tables (22) and (24), show that scores on the aptitude test correlate positively and significantly with the achievement test of Eng.120. The correlation coefficients are .520 and .297 for male and female students, respectively. Thus, on the basis of these results, the first null sub-hypothesis is rejected.

The second null sub-hypothesis states that there is no correlation between the students' scores on the cloze test and their scores on the achievement test of Eng.118, and between the same students' scores on the aptitude and the achievement test of Eng.118. Results show that scores on the cloze test correlate positively and significantly with scores on the achievement test of Eng.118. The correlation coefficients are .408 and .649 for the male and female students, respectively, as can be seen in tables (21) and (23). Results in tables (22) and (24) show that scores on the aptitude test correlate positively and significantly with the achievement test of Eng.118. The correlation coefficients are .501 and .389 with male and female students' scores, respectively. Thus, on the basis of these results, the second null sub-hypothesis is rejected.

The third null sub-hypothesis states that there is no correlation between the students' scores on the matching cloze test and the scores of

the same students on the end of semester Eng.113 achievement test, and between the same students' scores on the aptitude test and their scores on Eng.113 achievement test. Results show that scores of male and female students on the matching cloze test correlate positively and significantly with their scores on Eng.113 achievement test. The correlation coefficients are .519 and .714 with male and female students' scores, respectively, as can be seen in tables (21) and (23). Results in tables (22) and (24), show that scores on the aptitude test correlate positively and significantly with scores on Eng.113 achievement test. The correlation coefficients are .555 and .351 on the basis of male and female students' scores, respectively. Thus, the third null sub-hypothesis is rejected.

The fourth null sub-hypothesis states that there is no correlation between the students' scores on the matching cloze test and the scores of the same students on Eng.112 achievement test, and between scores of the same students on the aptitude test and their scores on Eng.112 achievement test. Results show that the matching cloze test correlates positively and significantly with scores on Eng.112 achievement test. The correlation coefficients are .638 and .550 with male and female students' scores, respectively, as can be seen in tables (21) and (23). Results in tables (22) and (24), reveal a positive and significant correlation between scores on the aptitude test and scores on Eng.112 achievement test, with the exception of the scores of female students on these variables. The

correlation between aptitude and Eng.112 scores of female students is very low and thus insignificant. Therefore, the fourth null sub-hypothesis is rejected considering the male group whereas it can not be rejected taking into account female students' scores, as it can be seen in tables (22) and (24).

The fifth null sub-hypothesis states that there is no correlation between the students' scores on the matching cloze test and the scores of the same students on Eng.111 achievement test, and between the same students' scores on the aptitude test and their scores on Eng.111 achievement test. Results show that the matching cloze test correlates positively and significantly with the students' scores on Eng.111 achievement test. The correlation coefficients are .387 and .435 with male and female students' scores, respectively. Moreover, results in tables (22) and (24) reveal that scores on the aptitude test correlate positively and significantly with scores on Eng.111 achievement test. The correlation coefficients are .447 and .387, considering male and female students' scores, respectively. Thus, the fifth null sub-hypothesis is rejected.

5.3 Implications of the Study

The results of this study have both theoretical and practical implications. Theoretically speaking, the results of the present study may provide answers to the questions of what underlying factors are involved

in matching cloze test performance. First, the relatively strong relationship between the matching cloze and the aptitude tests confirms the point that there is a close relationship between linguistic proficiency as measured by the matching cloze test and language aptitude as revealed by the aptitude test's scores. By extension, it could be claimed that one would expect increase in linguistic proficiency as aptitude increases.

Second, the highest correlation coefficient observed between Eng.113, the course of reading comprehension and the matching cloze test scores of female students for whom the cloze test was the best predictor, indicates that this procedure has potential for measuring integrative skills as they are evaluated by reading comprehension courses. Within the same context, it was quite understandable that the matching cloze test based on both groups' scores revealed to have the least predictive values with Eng.111, the course of basic language skills that seems to emphasize tasks that are not highly integrative in nature. To summarize, it could be stated that these remarks add to the theoretical implications of the matching cloze test which tend to correlate best with tests that require a certain level of integrative skills and aptitude. Moreover, they replicate the findings of many researchers such as Oller (1972), Turner (1989), Irvine, Atai and Oller (1974) and Alderson (1979).

Practically speaking, the results of this study may have several important implications. Considering the time constraints under which the

language teacher works and considering the large number of students admitted each semester, one should opt for the matching cloze test because of ease of administration and construction. However, it is not a good idea to develop a cloze test from any passage, since such a test may not function as a good test. The encouraging results of this study favors a careful selection of a passage that suites the reading level of the subjects to construct a cloze test. In this case, a carefully planned, developed, and piloted matching cloze test, which is both reliable and valid, should be adopted and used by the Department administrators and EFL teachers. Such use, we believe, has the potential for improving undergraduate instruction in these following ways: first, EFL teachers can make reasonable predictions about the probability of success of the students in the course.

Second, textbook selection can be improved by relying on the students' needs and reading abilities that are shown in their matching cloze test scores. These scores may provide powerful information about the degree of match or mismatch between the students' levels and their reading course materials.

Third, EFL teachers may use the information provided by the matching cloze test scores in making basic instructional decisions accordingly, speeding up on specific tasks, slowing down in others,

placing students in appropriate ability groups, and assigning appropriate and effective teaching strategies as required.

Moreover, it is suggested that similar academic systems which teach foreign languages may need to use the procedure of the matching cloze test in order to benefit from its predictive values.

In conclusion to this section, I would like to emphasize that although the present study is an important first step in investigating the predictive validity of the matching cloze test, the results should not be overgeneralized. It should be remembered that these findings and results are based upon a specific type of cloze procedure and upon a specific type of subjects, as described in this study. Therefore, the following section is devoted to provide some suggestions that may answer more questions about this field.

5.4 Suggestions for Further Research

This section provides several suggestions for more research in the same area to pursue in the future. These suggestions are given as follows:

1. It is suggested that the matching cloze test be longer in order to include more than 30 items, or it could have different deletion patterns, such as every fifth or sixth word deletions instead of the seventh word deletion used here.

2. It is also suggested to replicate the use of the matching cloze test as a predictive tool with other students in other academic institutions in the Kingdom of Saudi Arabia or with students in other Arab countries.
3. Unfortunately, the present study failed to use other forms of the cloze procedure with first year Saudi EFL students. Thus, a recommendation for further research would be to try out other forms of the cloze procedure such as the open-ended cloze or the multiple-choice cloze tests.
4. Since one level of learners was considered in the use of the matching cloze test, an area worth of investigation would be the inclusion of other levels such as the intermediate or advanced levels. For instance, within this principle of extension, students' scores on the matching cloze test could be correlated with their scores on their achievement tests administered at the end of second and third term of EFL undergraduate level.
5. Cloze procedure developers should also investigate and extend the use of the cloze procedure to predict not only EFL performance but also content area learning on which English is the medium of instruction. For example, the use of the matching cloze test as a predictive tool could be replicated with non-native students studying sociology or psychology in which English is the medium of instruction. Or, it could be replicated with students studying specific courses such as sociolinguistics, psycholinguistics, etc. to predict their performance in these areas

References

- Abraham, R. and Chapelle, C. (1992). The meaning of cloze test scores: an item difficulty perspective. *The Modern Language Journal*. Vol. 76, 468-479.
- Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly*. Vol. 13, No. 2, 219-227.
- Allen, Walter. (1980). Correlation of first day cloze with final grade in literature for foreign students. *Research in the Teaching of English*. Vol. 14, No. 1, 61-66.
- Ammeraal, Brenda. (1997). Correlation between students' placement scores on Acorn book multiple-choice tests and passing rate on the AP language examinations. (ERIC ED 403556).
- Ayers, J. and Peters, R. Martin. (1977). Predictive validity of test of English as a foreign language for Asian graduate students in engineering, chemistry or mathematics. *Educational and Psychological Measurement*. Vol. 37, No. 2, 461-463.
- Ayers, J. and Quattlebaum, R. (1992). TOEFL performance and success in Masters program in engineering. *Educational and Psychological Measurement*. Vol. 52, No. 4, 973-975.
- Bachman, Lyle. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*. Vol. 19, 535-556.
- Baldauf, R. and Propst, I. (1979). Measuring ESL reading achievement with matching cloze. Paper presented at the Micronesian Educators Conference, Pacific Islands (ED 186468).
- Bastidas, Jesus. (1984). The cloze procedure as a teaching technique to improve reading comprehension. *English Teaching Forum*. Vol. XXII, No. 2, 20-24.
- Bensoussan, M. and Ramraz, R. (1982). The fill-in test: combining multiple-choice and modified cloze techniques to test English as a foreign language reading comprehension. Report No. 52. Haifa: University Assessment and Selection Unit (ED 227652).

- Bostic, M. (1981). A correlational study of academic achievement and the test of English as a second language. *Dissertation Abstracts International*, Vol. 42, No. 20, 468-469.
- Brown, H. Douglas. (1994). *Principles of language learning and teaching* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall Regents.
- Brown J. D. (1993). What are the characteristics of natural cloze tests? *Language Testing*. Vol. 10, 93-115.
- Cohen, Andrew. (1980). *Testing language ability in the classroom*. Rowley, Massachusetts: Newbury House.
- _____. (1991). Second language testing. In Celce-Murica, Editor, *English as a second language*, USA, Newbury House.
- The College of the Canyons. "Predictive Validity Study of the APS Reading Test." (1994). Office of Institutional Development. (ERIC ED 374853).
- Cook, Vivian. (1996). *Second language learning and language teaching*. (2nd ed). London, New York: Arnold.
- Elaine, S., and others. (1986). The cloze procedure as predictor of undergraduate achievement in introductory courses. *Forum for Reading*. Vol. 18, No. 1, 32-36.
- Al-Fallay, Ibrahim. (1994). "The Impact of Background Knowledge on the Proficiency of Saudi Arabian Students Learning English as a Foreign Language." (Doctoral dissertation, The University of New Mexico).
- _____. (1997a). Investigating the reliability and validity of the fixed ratio multiple choice cloze test. *DIRASAT, An International Refereed Research Journal, Human and Social Sciences*. Vol. 24, No. 2, 507-526.
- _____. (1997b). The relationship between the cloze testing procedure and the skills tests of EFL: a correlational study. *Arab Journal for the Humanities*. No. 57, 347-3670

- Farhady, H. and M. N. Keramati. (1996). A text-driven method for the deletion procedure in cloze passages. *Language Testing*. Vol. 13, No. 2, 191-207.
- Flesch, Rudolf. (1948). A formula for predicting readability: instructions. *Journal of Applied Psychology*. Vol. 32, No. 3, 221-233.
- Fotos, Sandra. (1991). The cloze test as an integrative measure of EFL proficiency: a substitute for essays on college entrance examination. *Language Learning*. Vol. 41, No. 3, 313-336.
- Hanania, E. and M. Shikhani. (1986). Interrelationships among three tests of language proficiency: standardized ESL, cloze and writing. *TESOL Quarterly*, Vol. 20, No. 1, 97-109.
- Hanzeli, Victor. (1977). The effectiveness of cloze tests in measuring the competence of students of French in an academic setting. *The French Review*. Vol. L, No. 6, 865-874.
- Harris, David. (1969). *Testing English as a second language*. New York: McGraw-Hill.
- Heaton, J. B. (1988). *Writing English Language tests* (new ed.). London: Longman Group Limited.
- Hinofotis, F. (1980). Cloze as an alternative method of ESL Placement and Proficiency Testing. In Oller, J. W. jr and Perkins, K., editors, *Research in language testing*, Rowley, Mass.: Newbury House Publishers.
- Hisama, K. and others. (1977). A new direction in measuring proficiency in English as a second language. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY (ED 150198).
- Jonz, Jon. (1976). Improving on the basic egg: the M-C cloze. *Language Learning*. Vol. 26, No. 2, 255-265.
- Irvine, P., Atai, P. and Oller, John W. Jr. (1974). Cloze, dictation, and the test of English as a foreign language. *Language Learning*. Vol. 24, No. 2, 245-252.

- Al-Khalaf, Moodhy. (2000). "Predicting success of Saudi College FL Majors: an Adaptation of the Modern Language Aptitude Test (MLAT)." (MA dissertation, Girls' College of Education, Riyadh).
- Kubiszyn, T. and Borich, G. (1987). *Educational Testing and Measurement: Classroom Application and Practice* (2nd ed). Glenview, London: Scott, Foresman and Company.
- Lado, R. (1961). *Language testing: the construction and use of foreign language tests*. London: Longman, Green and Co. Ltd.
- Al-Musawi, N. and Al-Ansari, S. (1999). Test of English as a Foreign Language and First Certificate of English Tests as predictors of academic success for undergraduate students at the university of Bahrain. *System*. Vol. 27, No. 3, 389-399.
- Neal, Menka. (1998). The Predictive validity of the GRE and TOEFL exams with GGPA as the criterion of graduate success for international graduate students in science and engineering. (ERIC ED 424294).
- Norris, John. (2000). Purposeful language assessment: Selecting the right alternative test. *English Teaching Forum*. Vol. 38, No. 1, 18-23.
- Odunze, O. (1982). Test of English as a foreign language and first year GPA of Nigerian students. *Dissertation Abstracts International*, Vol. 42, No. 8, 3419-3420.
- Ojerinde, A. (1980). Cloze tests as a measure of English language proficiency among Nigerian primary-school pupils. *ELT Journal*. Vol. 15, 62-66.
- Oller, John W. Jr. (1972). Scoring methods and difficulty levels for cloze tests of proficiency in English as a second language. *The Modern Language Journal*. Vol. 56, 151-158.
- _____. (1973). Cloze tests of second language proficiency and what they measure. *Language Learning*. Vol. 23, No. 1, 105-118.

- _____.(1979). *Language tests at school: a pragmatic approach*. London: Longman Group Limited.
- Oller, J. W. jr and Perkins, K. (1978). *Language in Education: Testing the Tests*. Rowley, Mass.: Newbury House.
- _____.Ed. (1980). *Research in language testing*. Rowley, Mass.: Newbury House Publishers.
- Powell, J. and Gillespie, C. (1990). Assessment: all tests are not created equally. Paper presented at the Annual Meeting of the American Reading Forum, Sarasota, FL (ED 328908).
- Richards, J. and T. Rodgers. (1986). *Approaches and methods in language teaching: a description and analysis*. Cambridge: Cambridge University Press.
- Al-Salem, Reem. (1998). "The Predictive Validity of the English Admission Test Battery in the Department of European Languages and Translation." (MA dissertation, King Saud University, Riyadh).
- Sasaki, Miyuki. (2000). Effects of cultural schemata on students' test-taking processes for cloze tests: a multiple data source approach. *Language Testing*. Vol. 17, No. 1, 85-114.
- Schulz, Renate. (1981). Literature and readability: bridging the gap in foreign language reading. *Modern Language Journal*. Vol. 65, 43-53.
- Sharon, Amiel. (1971). Test of English as a Foreign Language as a Moderator of Graduate Record Examinations Scores in the Prediction of Foreign Students' Grades in Graduate School. (ERIC ED 058304).
- Soyoung, Lee. (1996). The concurrent validity of cloze test with essay test among Korean students. *Texas Papers in Foreign Language Education*. Vol. 2, No. 2, 57-69.
- Spiegel, Murray. (1991). *Shaum's outline of Theory and problems of statistics*. (2nd ed.). New York, St. Louis, London: McGRAW-HILL, INC.

- Spolsky, Bernard. (1986). A multiple choice for language testers. *Language Testing*. Vol. 3, No. 2, 147-158.
- Stansfield, Charles W. (1989). Language Aptitude Reconsidered. (ERIC ED 318226).
- Stephens, Elaine C. and others. (1986). The cloze procedure as predictor of undergraduate achievement in introductory courses. *Forum for Reading*. Vol. 18, No. 1, 32-36.
- Stubbs, J. and Tucker, R. (1974). The cloze test as a measure of ESL proficiency for Arab students. *The Modern Language Journal*. Vol. 58, 239-241.
- Turner, Carolyn. (1989). The underlying factor structure of L2 cloze test performance in francophone, university-level students: causal modeling as an approach to construct validation. *Language Testing*. Vol. 6, No. 2, 172-198.
- Valette, Rebecca. (1977). *Modern language testing: a handbook*. New York, Chicago, San Francisco, and Atlanta: Harcourt, Brace & World, INC.

Appendix 1

		:	
		:	
		:	
		:	
		:	
		:	-
	() ()	-	-
		-	-
		-	-
()	:()	:	-1
		:	-
()	()	:	-2
		:	-
()	()	:	-3
		:	-
	() ()	:	-
%		:	-
		:	-
)	-	-
	() () (-	-

Directions for taking this cloze test:

- 1- Read the whole passage at least once before you start to write. Note that the following passage is divided into eight sections.
- 2- Use the words in the boxes and only the words in the boxes next to each section to fill in the blanks. Each word should be used only once.
- 3- Provide only one word in each blank.
- 4- Use capital letters if the missing word comes at the beginning of the sentence.
- 5- Be careful! There is an extra word in each box for each section. That word does not fit in any blank.
- 6- Consider the following example:

Use the words in the box to fill in the blanks of the following sentence:

- The man left his and moved to a one.

The words that correctly fill in the blanks are “house” and “new”. The extra word that does not fit in any blank is “and”.

and
house
new

The cloze test

Language is how humans communicate with one another, either through speech, writing, or sign language. It comes from the Latin word lingua, meaning ‘tongue’.

Language started as a collection of, arranged so that they mean something thousands of years, people developed different and then organized the sounds into An alphabet is a collection of that represent the sounds we use.

alphabets
different
languages
over
sings
sounds

....., there are between 4,000 and 5,000 languages in World. Mandarin Chinese is the language by the largest number of people. come English, which is spoken in countries than any other language.

by
more
next
spoken
the
today

About 845 are spoken in India. People have to make up artificial universal languages. best known is Esperanto, invented in 1887, most widely used international language.

languages
the
the
tried
universal

Languages Change. New words such as 'Internet' into use. Old words disappear, for 'bombard', a canon. Latin, the language the Romans, is not spoken but can read it.

come
constantly
example
for
of
people

Words can also from one language to another. English many such words, such as 'video' Latin and 'planet' from Greek.

contains
from
many
move

Languages to families-groups of related languages developed from an original parent language Belongs to the Germanic branch of large Indo-European language family, while French to the Romance branch.

belong
belongs
English
the
to
which

Language is just confined to spoken or written Deaf people use sign language, while People read Braille, a special alphabet using raised dots. Even codes, such as those used by computers, are called languages.

blind
not
people
words

Appendix 2

Al-Khalaf designed an Arabic aptitude test to evaluate and measure the ability of Saudi students to learn a foreign language such as English. Al-Khalaf's aptitude test is composed of three parts, each dealing with a specific aspect. Part I consists of 50 items, part II of 45 items and part III includes 25 items. Following are examples of each part, an item analysis of part I and II, as well as a sample answer sheet all taken from Al-Khalaf (2000).

الجزء الأول : التعرف على الكلمات

] - () _
[
 .
 :
 :

[" " ":()]	: () () () () ()
---------------	-----------------------

" " " = ":() [: () () () () ()
-----------------	-----------------------

" " " ":() [" " " "	: () () () () ()
----------------------	-----------------------

الجزء الثاني: وظائف الكلمات في الجمل

:

:

() () () () ()

" " " " " " " "

:

() () () () ()

" " " " " " " "

(.)

-

()

-

()

-

() () ()

" " " " " "

" " " "

-

-

!

الجزء الثالث : الأزواج المترادفة

_____ :

()

_____ - _____	_____ - _____
---------------	---------------

:

:" "

[()

" "]

()

()

()

()

()

!

Part I: Item analysis

Item No.	Index of Popularity	Index of Difficulty	Index of Discrimination
1	0.65	0.35	0.22
2	0.75	0.25	0.18
3	0.80	0.20	0.16
4	0.65	0.35	0.22
5	0.70	0.30	0.21
6	0.60	0.40	0.24
7	0.65	0.35	0.22
8	0.40	0.60	0.24
9	0.80	0.20	0.16
10	0.55	0.45	0.24
11	0.60	0.40	0.24
12	0.75	0.25	0.18
13	0.75	0.25	0.18
14	0.26	0.75	0.18
15	0.75	0.25	0.18
16	0.60	0.40	0.24
17	0.45	0.55	0.24
18	0.70	0.30	0.21
19	0.75	0.25	0.18
20	0.50	0.50	0.25
21	0.70	0.30	0.21
22	0.20	0.80	0.16
23	0.75	0.25	0.18
24	0.70	0.30	0.21
25	0.45	0.55	0.24
26	0.80	0.20	0.16
27	0.45	0.55	0.24
28	0.45	0.55	0.24
29	0.60	0.40	0.24
30	0.20	0.80	0.16
31	0.60	0.40	0.24
32	0.75	0.25	0.18
33	0.25	0.75	0.18
34	0.50	0.50	0.25
35	0.60	0.40	0.24
36	0.20	0.80	0.16
37	0.30	0.70	0.21
38	0.40	0.60	0.24
39	0.30	0.70	0.21
40	0.25	0.75	0.18
41	0.30	0.70	0.21
42	0.50	0.50	0.25
43	0.20	0.80	0.16
44	0.20	0.80	0.16
45	0.30	0.70	0.21

46	0.25	0.75	0.18
47	0.20	0.80	0.16
48	0.25	0.75	0.18
49	0.35	0.65	0.22
50	0.20	0.80	0.16

Part II: Item analysis

Item No.	Index of Popularity	Index of Difficulty	Index of Discrimination
1	0.70	0.30	0.21
2	0.25	0.75	0.18
3	0.55	0.45	0.24
4	0.30	0.70	0.21
5	0.65	0.35	0.22
6	0.25	0.75	0.18
7	0.65	0.35	0.22
8	0.40	0.60	0.24
9	0.50	0.50	0.25
10	0.60	0.40	0.24
11	0.45	0.55	0.24
12	0.75	0.25	0.18
13	0.35	0.65	0.22
14	0.20	0.80	0.16
15	0.35	0.65	0.22
16	0.80	0.20	0.16
17	0.45	0.55	0.24
18	0.65	0.35	0.22
19	0.70	0.30	0.21
20	0.25	0.75	0.18
21	0.80	0.20	0.16
22	0.80	0.20	0.16
23	0.35	0.65	0.22
24	0.45	0.55	0.24
25	0.20	0.80	0.16
26	0.30	0.70	0.21
27	0.65	0.35	0.22
28	0.80	0.20	0.16
29	0.75	0.25	0.18
30	0.50	0.50	0.25
31	0.45	0.55	0.24
32	0.25	0.75	0.18
33	0.65	0.35	0.22
34	0.65	0.35	0.22
35	0.80	0.20	0.16
36	0.75	0.25	0.18
37	0.35	0.65	0.22
38	0.25	0.75	0.18
39	0.25	0.75	0.18
40	0.45	0.55	0.24
41	0.60	0.40	0.24
42	0.50	0.50	0.25

43	0.55	0.45	0.24
44	0.20	0.80	0.16
45	0.50	0.50	0.25

اختبار الاستعداد للتعلم اللغوي الحديث

() :
/ :

الجزء الثاني (وظائف الكلمات في جمل)				
الأمثلة: (ن) أ ب ج د هـ				
(هـ)	أ	ب	ج	د
-1	أ	ب	ج	د
-2	أ	ب	ج	د
-3	أ	ب	ج	د
-4	أ	ب	ج	د
-5	أ	ب	ج	د
-6	أ	ب	ج	د
-7	أ	ب	ج	د
-8	أ	ب	ج	د
-9	أ	ب	ج	د
-10	أ	ب	ج	د
-11	أ	ب	ج	د
-12	أ	ب	ج	د
-13	أ	ب	ج	د
-14	أ	ب	ج	د
-15	أ	ب	ج	د
-16	أ	ب	ج	د
-17	أ	ب	ج	د
-18	أ	ب	ج	د
-19	أ	ب	ج	د
-20	أ	ب	ج	د
-21	أ	ب	ج	د
-22	أ	ب	ج	د
-23	أ	ب	ج	د
-24	أ	ب	ج	د
-25	أ	ب	ج	د
-26	أ	ب	ج	د
-27	أ	ب	ج	د
-28	أ	ب	ج	د
-29	أ	ب	ج	د
-30	أ	ب	ج	د
-31	أ	ب	ج	د
-32	أ	ب	ج	د
-33	أ	ب	ج	د
-34	أ	ب	ج	د
-35	أ	ب	ج	د
-36	أ	ب	ج	د
-37	أ	ب	ج	د
-38	أ	ب	ج	د
-39	أ	ب	ج	د
-40	أ	ب	ج	د
-41	أ	ب	ج	د
-42	أ	ب	ج	د
-43	أ	ب	ج	د
-44	أ	ب	ج	د
-45	أ	ب	ج	د

الجزء الأول (التعرف على الكلمات)				
الأمثلة: (ك) أ ب ج د هـ				
(ل)	أ	ب	ج	د
-1	أ	ب	ج	د
-2	أ	ب	ج	د
-3	أ	ب	ج	د
-4	أ	ب	ج	د
-5	أ	ب	ج	د
-6	أ	ب	ج	د
-7	أ	ب	ج	د
-8	أ	ب	ج	د
-9	أ	ب	ج	د
-10	أ	ب	ج	د
-11	أ	ب	ج	د
-12	أ	ب	ج	د
-13	أ	ب	ج	د
-14	أ	ب	ج	د
-15	أ	ب	ج	د
-16	أ	ب	ج	د
-17	أ	ب	ج	د
-18	أ	ب	ج	د
-19	أ	ب	ج	د
-20	أ	ب	ج	د
-21	أ	ب	ج	د
-22	أ	ب	ج	د
-23	أ	ب	ج	د
-24	أ	ب	ج	د
-25	أ	ب	ج	د
-26	أ	ب	ج	د
-27	أ	ب	ج	د
-28	أ	ب	ج	د
-29	أ	ب	ج	د
-30	أ	ب	ج	د
-31	أ	ب	ج	د
-32	أ	ب	ج	د
-33	أ	ب	ج	د
-34	أ	ب	ج	د
-35	أ	ب	ج	د
-36	أ	ب	ج	د
-37	أ	ب	ج	د
-38	أ	ب	ج	د
-39	أ	ب	ج	د
-40	أ	ب	ج	د
-41	أ	ب	ج	د
-42	أ	ب	ج	د
-43	أ	ب	ج	د
-44	أ	ب	ج	د
-45	أ	ب	ج	د
-46	أ	ب	ج	د
-47	أ	ب	ج	د
-48	أ	ب	ج	د
-49	أ	ب	ج	د
-50	أ	ب	ج	د



الجزء الأول (التعرف على الكلمات)					المثال: (ي)
هـ	د	ج	ب	أ	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-1
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-2
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-3
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-4
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-6
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-7
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-8
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-9
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-10
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-11
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-12
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-13
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-14
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-15
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-16
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-17
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-18
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-19
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-20
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-21
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-22
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-23
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-24
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-25

